

μL of the diluted conjugate was added to each well and incubated for 30 minutes at room temperature. Following incubation, the wells were washed five times with PBS/0.1% Tween 20™. 100 μL of tetramethylbenzidine peroxidase (TMB) substrate (Kirkegaard and Perry Laboratories, Gaithersburg, MD) was added, undiluted, and incubated for about 15 minutes. The reaction was stopped with the addition of 100 μL of 1 N H<sub>2</sub>SO<sub>4</sub> to each well, and the plates were read at 450 nm.

Figure 4 shows the ELISA reactivity of two recombinant antigens isolated using method A in Example 3 (TbRa3 and TbRa9) with sera from *M. tuberculosis* positive and negative patients. The reactivity of these antigens is compared to that of bacterial lysate isolated from *M. tuberculosis* strain H37Ra (Difco, Detroit, MI). In both cases, the recombinant antigens differentiated positive from negative sera. Based on cut-off values obtained from receiver-operator curves, TbRa3 detected 56 out of 87 positive sera, and TbRa9 detected 111 out of 165 positive sera.

Figure 5 illustrates the ELISA reactivity of representative antigens isolated using method B of Example 3. The reactivity of the recombinant antigens TbH4, TbH12, Tb38-1 and the peptide TbM-1 (as described in Example 4) is compared to that of the 38 kD antigen described by Andersen and Hansen, *Infect. Immun.* 57:2481-2488, 1989. Again, all of the polypeptides tested differentiated positive from negative sera. Based on cut-off values obtained from receiver-operator curves, TbH4 detected 67 out of 126 positive sera, TbH12 detected 50 out of 125 positive sera, 38-1 detected 61 out of 101 positive sera and the TbM-1 peptide detected 25 out of 30 positive sera.

The reactivity of four antigens (TbRa3, TbRa9, TbH4 and TbH12) with sera from a group of *M. tuberculosis* infected patients with differing reactivity in the acid fast stain of sputum (Smithwick and David, *Tubercle* 52:226, 1971) was also examined, and compared to the reactivity of *M. tuberculosis* lysate and the 38 kD antigen. The results are presented in Table 3, below:

TABLE 3

REACTIVITY OF ANTIGENS WITH SERA FROM *M. TUBERCULOSIS* PATIENTS

Patient	Acid Fast Sputum	ELISA Values					
		Lysate	38kD	TbRa9	TbH12	TbH4	TbRa3
Tb01B93I-2	+++	1.853	0.634	0.998	1.022	1.030	1.314
Tb01B93I-19	+++	2.657	2.322	0.608	0.837	1.857	2.335
Tb01B93I-8	+++	2.703	0.527	0.492	0.281	0.501	2.002
Tb01B93I-10	++	1.665	1.501	0.685	0.216	0.448	0.458
Tb01B93I-11	+++	2.817	0.697	0.509	0.301	0.173	2.608
Tb01B93I-15	+++	1.28	0.283	0.808	0.218	1.537	0.811
Tb01B93I-16	+++	2.908	>3	0.899	0.441	0.593	1.080
Tb01B93I-25	---	0.395	0.131	0.335	0.211	0.107	0.948
Tb01B93I-87	+++	2.653	2.432	2.282	0.977	1.221	0.857
Tb01B93I-89	---	1.912	2.370	2.436	0.876	0.520	0.952
Tb01B94I-108	---	1.639	0.341	0.797	0.368	0.654	0.798
Tb01B94I-201	---	1.721	0.419	0.661	0.137	0.064	0.692
Tb01B93I-88	---	1.939	1.269	2.519	1.381	0.214	0.530
Tb01B93I-92	---	2.555	2.329	2.78	0.685	0.997	2.527
Tb01B94I-109	---	0.993	0.620	0.574	0.441	0.5	2.558
Tb01B94I-210	++	2.777	>3	0.393	0.367	1.004	1.315
Tb01B94I-224	++	2.913	0.476	0.251	1.297	1.990	0.256
Tb01B93I-9	+	2.649	0.278	0.210	0.140	0.181	1.586
Tb01B93I-14	+	>3	1.538	0.282	0.291	0.549	2.880
Tb01B93I-21	+	2.645	0.739	2.499	0.783	0.536	1.770

Patient	Acid Fast Sputum	ELISA Values					
		Lysate	38kD	TbRa9	TbH12	TbH4	TbRa3
Tb01B93I-22	+	0.714	0.451	2.082	0.285	0.269	1.159
Tb01B93I-31	+	0.956	0.490	1.019	0.812	0.176	1.293
Tb01B93I-32	-	2.261	0.786	0.668	0.273	0.535	0.405
Tb01B93I-52	-	0.658	0.114	0.434	0.330	0.273	1.140
Tb01B93I-99	-	2.118	0.584	1.62	0.119	0.977	0.729
Tb01B94I-130	-	1.349	0.224	0.86	0.282	0.383	2.146
Tb01B94I-131	-	0.685	0.324	1.173	0.059	0.118	1.431
AT4-0070	Normal	0.072	0.043	0.092	0.071	0.040	0.039
AT4-0105	Normal	0.397	0.121	0.118	0.103	0.078	0.390
3/15/94-1	Normal	0.227	0.064	0.098	0.026	0.001	0.228
4/15/93-2	Normal	0.114	0.240	0.071	0.034	0.041	0.264
5/26/94-4	Normal	0.089	0.259	0.096	0.046	0.008	0.053
5/26/94-3	Normal	0.139	0.093	0.085	0.019	0.067	0.01

Based on cut-off values obtained from receiver-operator curves, TbRa3 detected 23 out of 27 positive sera, TbRa9 detected 22 out of 27, TbH4 detected 18 out of 27 and TbH12 detected 15 out of 27. If used in combination, these four antigens would have a theoretical sensitivity of 27 out of 27, indicating that these antigens should complement each other in the serological detection of *M. tuberculosis* infection. In addition, several of the recombinant antigens detected positive sera that were not detected using the 38 kD antigen, indicating that these antigens may be complementary to the 38 kD antigen.

The reactivity of the recombinant antigen TbRa11 with sera from *M. tuberculosis* patients shown to be negative for the 38 kD antigen, as well as with sera from PPD positive and normal donors, was determined by ELISA as described above.

The results are shown in Figure 6 which indicates that TbRa11, while being negative with sera from PPD positive and normal donors, detected sera that were negative with the 38 kD antigen. Of the thirteen 38 kD negative sera tested, nine were positive with TbRa11, indicating that this antigen may be reacting with a sub-group of 38 kD antigen negative sera. In contrast, in a group of 38 kD positive sera where TbRa11 was reactive, the mean OD 450 for TbRa11 was lower than that for the 38 kD antigen. The data indicate an inverse relationship between the presence of TbRa11 activity and 38 kD positivity.

The antigen TbRa2A was tested in an indirect ELISA using initially 50  $\mu$ l of serum at 1:100 dilution for 30 minutes at room temperature followed by washing in PBS Tween and incubating for 30 minutes with biotinylated Protein A (Zymed, San Francisco, CA) at a 1:10,000 dilution. Following washing, 50  $\mu$ l of streptavidin-horseradish peroxidase (Zymed) at 1:10,000 dilution was added and the mixture incubated for 30 minutes. After washing, the assay was developed with TMB substrate as described above. The reactivity of TbRa2A with sera from *M. tuberculosis* patients and normal donors is shown in Table 4. The mean value for reactivity of TbRa2A with sera from *M. tuberculosis* patients was 0.444 with a standard deviation of 0.309. The mean for reactivity with sera from normal donors was 0.109 with a standard deviation of 0.029. Testing of 38 kD negative sera (Figure 7) also indicated that the TbRa2A antigen was capable of detecting sera in this category.

TABLE 4  
REACTIVITY OF TBRA2A WITH SERA FROM *M. TUBERCULOSIS* PATIENTS AND FROM  
NORMAL DONORS

Serum ID	Status	OD 450
Tb85	TB	0.680
Tb86	TB	0.450
Tb87	TB	0.263
Tb88	TB	0.275
Tb89	TB	0.403
Tb91	TB	0.393
Tb92	TB	0.401

Tb93	TB	0.232
Tb94	TB	0.333
Tb95	TB	0.435
Tb96	TB	0.284
Tb97	TB	0.320
Tb99	TB	0.328
Tb100	TB	0.817
Tb101	TB	0.607
Tb102	TB	0.191
Tb103	TB	0.228
Tb107	TB	0.324
Tb109	TB	1.572
Tb112	TB	0.338
DL4-0176	Normal	0.036
AT4-0043	Normal	0.126
AT4-0044	Normal	0.130
AT4-0052	Normal	0.135
AT4-0053	Normal	0.133
AT4-0062	Normal	0.128
AT4-0070	Normal	0.088
AT4-0091	Normal	0.108
AT4-0100	Normal	0.106
AT4-0105	Normal	0.108
AT4-0109	Normal	0.105

The reactivity of the recombinant antigen (g) (SEQ ID NO: 60) with sera from *M. tuberculosis* patients and normal donors was determined by ELISA as described above. Figure 8 shows the results of the titration of antigen (g) with four *M. tuberculosis* positive sera that were all reactive with the 38 kD antigen and with four donor sera. All four positive sera were reactive with antigen (g).

The reactivity of the recombinant antigen TbH-29 (SEQ ID NO: 137) with sera from *M. tuberculosis* patients, PPD positive donors and normal donors was determined by indirect ELISA as described above. The results are shown in Figure 9. TbH-29 detected 30 out of 60 *M. tuberculosis* sera, 2 out of 8 PPD positive sera and 2 out of 27 normal sera.

Figure 10 shows the results of ELISA tests (both direct and indirect) of the antigen TbH-33 (SEQ ID NO: 140) with sera from *M. tuberculosis* patients and from normal donors and with a pool of sera from *M. tuberculosis* patients. The mean

OD 450 was demonstrated to be higher with sera from *M. tuberculosis* patients than from normal donors, with the mean OD 450 being significantly higher in the indirect ELISA than in the direct ELISA. Figure 11 is a titration curve for the reactivity of recombinant TbH-33 with sera from *M. tuberculosis* patients and from normal donors  
5 showing an increase in OD 450 with increasing concentration of antigen.

The reactivity of the recombinant antigens RDIF6, RDIF8 and RDIF10 (SEQ ID NOS: 184-187, respectively) with sera from *M. tuberculosis* patients and normal donors was determined by ELISA as described above. RDIF6 detected 6 out of 32 *M. tuberculosis* sera and 0 out of 15 normal sera; RDIF8 detected 14 out of 32 *M. tuberculosis* sera and 0 out of 15 normal sera; and RDIF10 detected 4 out of 27 *M. tuberculosis* sera and 1 out of 15 normal sera. In addition, RDIF10 was found to detect  
10 0 out of 5 sera from PPD-positive donors.

The antigens MO-1, MO-2, MO-4, MO-28 and MO-29 described above in Example 5, were expressed in *E. coli* and purified using a hexahistidine tag. The  
15 reactivity of these antigens with both *M. tuberculosis* positive and negative sera was examined by ELISA as described above. Titration curves showing the reactivity of MO-1, MO-2, MO-4, MO-28 and MO-29 at different solid phase coat levels when tested against four *M. tuberculosis* positive sera and four *M. tuberculosis* negative sera are shown in Figs. 12A-E, respectively. Three of the clones, MO-1, MO-2 and MO-29  
20 were further tested on panels of HIV positive/tuberculosis (HIV/TB) positive and extrapulmonary sera. MO-1 detected 3/20 extrapulmonary and 2/38 HIV/TB sera. On the same sera groups, MO-2 detected 2/20 and 10/38, and MO-29 detected 2/20 and 8/38 sera. In combination these three clones would have detected 4/20 extrapulmonary sera and 16/38 HIV/TB sera. In addition, MO-1 detected 6/17 sera that had previously  
25 been shown only to react with *M. tuberculosis* lysate and not with either 38 kD or with other antigens of the subject invention.

### EXAMPLE 10

#### PREPARATION AND CHARACTERIZATION OF *M. TUBERCULOSIS* FUSION PROTEINS

A fusion protein containing TbRa3, the 38 kD antigen and Tb38-1 was  
5 prepared as follows.

Each of the DNA constructs TbRa3, 38 kD and Tb38-1 were modified  
by PCR in order to facilitate their fusion and the subsequent expression of the fusion  
protein TbRa3-38 kD-Tb38-1. TbRa3, 38 kD and Tb38-1 DNA was used to perform  
PCR using the primers PDM-64 and PDM-65 (SEQ ID NO: 141 and 142), PDM-57 and  
10 PDM-58 (SEQ ID NO: 143 and 144), and PDM-69 and PDM-60 (SEQ ID NO: 145-  
146), respectively. In each case, the DNA amplification was performed using 10 µl  
10X Pfu buffer, 2 µl 10 mM dNTPs, 2 µl each of the PCR primers at 10 µM  
concentration, 81.5 µl water, 1.5 µl Pfu DNA polymerase (Stratagene, La Jolla, CA)  
and 1 µl DNA at either 70 ng/µl (for TbRa3) or 50 ng/µl (for 38 kD and Tb38-1). For  
15 TbRa3, denaturation at 94°C was performed for 2 min, followed by 40 cycles of 96°C  
for 15 sec and 72°C for 1 min, and lastly by 72°C for 4 min. For 38 kD, denaturation at  
96°C was performed for 2 min, followed by 40 cycles of 96°C for 30 sec, 68°C for 15  
sec and 72°C for 3 min, and finally by 72°C for 4 min. For Tb38-1 denaturation at 94°  
C for 2 min was followed by 10 cycles of 96°C for 15 sec, 68°C for 15 sec and 72°C for  
20 1.5 min, 30 cycles of 96°C for 15 sec, 64°C for 15 sec and 72°C for 1.5, and finally by  
72°C for 4 min.

The TbRa3 PCR fragment was digested with NdeI and EcoRI and cloned  
directly into pT7<sup>+</sup>L2 IL 1 vector using NdeI and EcoRI sites. The 38 kD PCR fragment  
was digested with Sse8387I, treated with T4 DNA polymerase to make blunt ends and  
25 then digested with EcoRI for direct cloning into the pT7<sup>+</sup>L2Ra3-1 vector which was  
digested with StuI and EcoRI. The 38-1 PCR fragment was digested with Eco47III and  
EcoRI and directly subcloned into pT7<sup>+</sup>L2Ra3/38kD-17 digested with the same  
enzymes. The whole fusion was then transferred to pET28b using NdeI and EcoRI  
sites. The fusion construct was confirmed by DNA sequencing.

The expression construct was transformed to BLR pLys S *E. coli* (Novagen, Madison, WI) and grown overnight in LB broth with kanamycin (30 µg/ml) and chloramphenicol (34 µg/ml). This culture (12 ml) was used to inoculate 500 ml 2XYT with the same antibiotics and the culture was induced with IPTG at an OD<sub>560</sub> of 0.44 to a final concentration of 1.2 mM. Four hours post-induction, the bacteria were harvested and sonicated in 20 mM Tris (8.0), 100 mM NaCl, 0.1% DOC, 20 µg/ml Leupeptin, 20 mM PMSF followed by centrifugation at 26,000 X g. The resulting pellet was resuspended in 8 M urea, 20 mM Tris (8.0), 100 mM NaCl and bound to Pro-bond nickel resin (Invitrogen, Carlsbad, CA). The column was washed several times with the above buffer then eluted with an imidazole gradient (50 mM, 100 mM, 500 mM imidazole was added to 8 M urea, 20 mM Tris (8.0), 100 mM NaCl). The eluates containing the protein of interest were then dialyzed against 10 mM Tris (8.0).

The DNA and amino acid sequences for the resulting fusion protein (hereinafter referred to as TbRa3-38 kD-Tb38-1) are provided in SEQ ID NO: 147 and 148, respectively.

A fusion protein containing the two antigens TbH9 and Tb38-1 (hereinafter referred to as TbH9-Tb38-1) without a hinge sequence, was prepared using a similar procedure to that described above. The DNA sequence for the TbH9-Tb38-1 fusion protein is provided in SEQ ID NO: 151.

A fusion protein containing TbRa3, the antigen 38kD, Tb38-1 and DPEP was prepared as follows.

Each of the DNA constructs TbRa3, 38 kD and Tb38-1 were modified by PCR and cloned into vectors essentially as described above, with the primers PDM-69 (SEQ ID NO:145 and PDM-83 (SEQ ID NO: 200) being used for amplification of the Tb38-1A fragment. Tb38-1A differs from Tb38-1 by a DraI site at the 3' end of the coding region that keeps the final amino acid intact while creating a blunt restriction site that is in frame. The TbRa3/38kD/Tb38-1A fusion was then transferred to pET28b using NdeI and EcoRI sites.

DPEP DNA was used to perform PCR using the primers PDM-84 and PDM-85 (SEQ ID NO: 201 and 202, respectively) and 1 µl DNA at 50 ng/µl.



Denaturation at 94 °C was performed for 2 min, followed by 10 cycles of 96 °C for 15 sec, 68 °C for 15 sec and 72 °C for 1.5 min; 30 cycles of 96 °C for 15 sec, 64 °C for 15 sec and 72 °C for 1.5 min; and finally by 72 °C for 4 min. The DPEP PCR fragment was digested with EcoRI and Eco72I and clones directly into the pET28Ra3/38kD/38-1A construct which was digested with DraI and EcoRI. The fusion construct was confirmed to be correct by DNA sequencing. Recombinant protein was prepared as described above. The DNA and amino acid sequences for the resulting fusion protein (hereinafter referred to as TbF-2) are provided in SEQ ID NO: 203 and 204, respectively.

A fusion protein containing TbRa3, the antigen 38kD, Tb38-1 and TbH4 was prepared as follows.

Genomic *M. tuberculosis* DNA was used to PCR full-length TbH4 (FL TbH4) with the primers PDM-157 and PDM-160 (SEQ ID NO: 343 and 344, respectively) and 2 µl DNA at 100 ng/µl. Denaturation at 96 °C was performed for 2 min, followed by 40 cycles of 96 °C for 30 sec, 61 °C for 20 sec and 72 °C for 5 min; and finally by annealing at 72 °C for 10 min. The FL TbH4 PCR fragment was digested with EcoRI and Sca I (New England Biolabs.) and cloned directly into the pET28Ra3/38kD/38-1A construct described above which was digested with DraI and EcoRI. The fusion construct was confirmed to be correct by DNA sequencing. Recombinant protein was prepared as described above. The DNA and amino acid sequences for the resulting fusion protein (hereinafter referred to as TbF-6) are provided in SEQ ID NO: 345 and 346, respectively.

A fusion protein containing the antigen 38kD and DPEP separated by a linker was prepared as follows.

38 kD DNA was used to perform PCR using the primers PDM-176 and PDM-175 (SEQ ID NO: 347 and 348, respectively), and 1 µl PET28Ra3/38kD/38-1/Ra2A-12 DNA at 110 ng/µl. Denaturation at 96 °C was performed for 2 min, followed by 40 cycles of 96 °C for 30 sec, 71 °C for 15 sec and 72 °C for 5 min and 40 sec; and finally by annealing at 72 °C for 4 min. The two sets of primers PDM-171, PDM-172, and PDM-173, PDM-174 were annealed by heating to 95 °C for 2 min and

then ramping down to 25 °C slowly at 0.1 °C/sec. DPEP DNA was used to perform PCR as described above. The 38 kD fragment was digested with Eco RI (New England Biolabs) and cloned into a modified pT7AL2 vector which was cut with Eco 72 I (Promega) and Eco RI. The modified pT7AL2 construct was designed to have a MGHHHHHH amino acid coding region in frame just 5' of the Eco 72 I site. The construct was digested with Kpn 2I (Gibco, BRL) and Pst I (New England Biolabs) and the annealed sets of phosphorylated primers (PDM-171, PDM-172 and PDM-173, PDM-174) were cloned in. The DPEP PCR fragment was digested with Eco RI and Eco 72 I and cloned into this second construct which was digested with Eco 47 III (New England Biolabs) and Eco RI. Ligations were done with a ligation kit from Panvera (Madison, WI). The resulting construct was digested with NdeI (New England Biolabs) and Eco RI, and transferred to a modified pET28 vector. The fusion construct was confirmed to be correct by DNA sequencing.

Recombinant protein was prepared essentially as described above. The DNA and amino acid sequences for the resulting fusion protein (hereinafter referred to as TbF-8) are provided in SEQ ID NO: 349 and 350, respectively.

## EXAMPLE II

### USE OF *M. TUBERCULOSIS* FUSION PROTEINS FOR SERODIAGNOSIS OF TUBERCULOSIS

The effectiveness of the fusion protein TbRa3-38 kD-Tb38-1, prepared as described above, in the serodiagnosis of tuberculosis infection was examined by ELISA.

The ELISA protocol was as described above in Example 6, with the fusion protein being coated at 200 ng/well. A panel of sera was chosen from a group of tuberculosis patients previously shown, either by ELISA or by western blot analysis, to react with each of the three antigens individually or in combination. Such a panel enabled the dissection of the serological reactivity of the fusion protein to determine if

all three epitopes functioned with the fusion protein. As shown in Table 5, all four sera that reacted with TbRa3 only were detectable with the fusion protein. Three sera that reacted only with Tb38-1 were also detectable, as were two sera that reacted with 38 kD alone. The remaining 15 sera were all positive with the fusion protein based on a cut-off in the assay of mean negatives +3 standard deviations. This data demonstrates the functional activity of all three epitopes in the fusion protein.

TABLE 5  
REACTIVITY OF TRI-PEPTIDE FUSION PROTEIN WITH SERA FROM *M. TUBERCULOSIS* PATIENTS

Serum ID	Status	ELISA and/or Western Blot Reactivity with Individual proteins			Fusion Recombinant OD 450	Fusion Recombinant Status
		38kd	Tb38-1	TbRa3		
01B931-40	TB	-	-	+	0.413	+
01B931-41	TB	-	+	+	0.392	+
01B931-29	TB	+	-	+	2.217	+
01B931-109	TB	+	±	+	0.522	+
01B931-132	TB	-	+	-	0.937	+
5004	TB	±	-	±	1.098	+
15004	TB	-	-	-	2.077	+
39004	TB	-	+	-	1.675	+
68004	TB	+	±	±	2.388	+
99004	TB	-	+	±	0.607	+
107004	TB	-	-	±	0.667	+
92004	TB	-	±	±	1.070	+
97004	TB	+	-	±	1.152	+
118004	TB	-	-	±	2.694	+
173004	TB	+	+	+	3.258	+
175004	TB	+	-	+	2.514	+
274004	TB	-	-	±	3.220	+
276004	TB	-	+	-	2.991	+
282004	TB	-	-	-	0.824	+
289004	TB	-	-	-	0.848	+

308004	TB	-	+	-	3.338	+
314004	TB	-	+	-	1.362	+
317004	TB	+	-	-	0.763	+
312004	TB	-	-	+	1.079	+
D176	PPD	-	-	-	0.145	-
D162	PPD	-	-	-	0.073	-
D161	PPD	-	-	-	0.097	-
D27	PPD	-	-	-	0.082	-
A6-124	NORMAL	-	-	-	0.053	-
A6-125	NORMAL	-	-	-	0.087	-
A6-126	NORMAL	-	-	-	0.346	-
A6-127	NORMAL	-	-	-	0.064	-
A6-128	NORMAL	-	-	-	0.034	-
A6-129	NORMAL	-	-	-	0.037	-
A6-130	NORMAL	-	-	-	0.057	-
A6-131	NORMAL	-	-	-	0.054	-
A6-132	NORMAL	-	-	-	0.022	-
A6-133	NORMAL	-	-	-	0.147	-
A6-134	NORMAL	-	-	-	0.101	-
A6-135	NORMAL	-	-	-	0.066	-
A6-136	NORMAL	-	-	-	0.054	-
A6-137	NORMAL	-	-	-	0.063	-
A6-138	NORMAL	-	-	-	0.041	-
A6-139	NORMAL	-	-	-	0.103	-
A6-140	NORMAL	-	-	-	0.212	-
A6-141	NORMAL	-	-	-	0.056	-
A6-142	NORMAL	-	-	-	0.051	-

The reactivity of the fusion protein TbF-2 with sera from *M. tuberculosis*-infected patients was examined by ELISA using the protocol described above. The results of these studies (Table 6) demonstrate that all four antigens function independently in the fusion protein.

TABLE 6  
 REACTIVITY OF TbF-2 FUSION PROTEIN WITH TB AND NORMAL SERA

Serum ID	Status	TbF OD450	Status	TbF-2 OD450	Status	ELISA Reactivity			
						3B ID	TbRaJ	Tb35-1	DPEP
B931-40	TB	0.57	-	0.321	-	-	-	-	-
B931-41	TB	0.601	-	0.396	-	-	-	-	-
B931-109	TB	0.494	-	0.404	-	-	-	-	-
B931-132	TB	1.502	-	1.292	-	-	-	2	-
5804	TB	1.806	-	1.666	-	-	-	-	-
15004	TB	2.862	-	2.468	-	-	-	-	-
39004	TB	2.443	-	1.722	-	-	-	-	-
68004	TB	2.871	-	2.575	-	-	-	-	-
99004	TB	0.691	-	0.971	-	-	-	-	-
107004	TB	0.575	-	0.732	-	-	-	-	-
92004	TB	1.632	-	1.394	-	-	-	-	-
97004	TB	1.491	-	1.979	-	-	-	2	-
118004	TB	3.182	-	3.045	-	-	-	-	-
173004	TB	2.644	-	2.578	-	-	-	-	-
175004	TB	3.232	-	2.916	-	-	-	-	-
274004	TB	2.696	-	3.716	-	-	-	-	-
276004	TB	3.243	-	2.56	-	-	-	-	-
282004	TB	1.249	-	1.234	-	-	-	-	-
289004	TB	1.373	-	1.17	-	-	-	-	-
508004	TB	3.708	-	3.353	-	-	-	-	-
714004	TB	1.663	-	1.399	-	-	-	-	-
717004	TB	1.162	-	0.92	-	-	-	-	-
712004	TB	1.709	-	1.453	-	-	-	-	-
880004	TB	0.278	-	0.461	-	-	-	-	-
451004	TB	0.18	-	0.2	-	-	-	-	-
478004	TB	0.188	-	0.469	-	-	-	-	-
410004	TB	0.364	-	2.392	-	-	-	-	-
411004	TB	0.306	-	0.874	-	-	-	-	-
421004	TB	0.357	-	1.456	-	-	-	-	-
528004	TB	0.947	-	0.196	-	-	-	-	-
A6-87	Normal	0.094	-	0.063	-	-	-	-	-
A6-88	Normal	0.214	-	0.19	-	-	-	-	-
A6-89	Normal	0.248	-	0.123	-	-	-	-	-
A6-90	Normal	0.179	-	0.206	-	-	-	-	-
A6-91	Normal	0.133	-	0.151	-	-	-	-	-
A6-92	Normal	0.064	-	0.097	-	-	-	-	-
A6-93	Normal	0.072	-	0.098	-	-	-	-	-
A6-94	Normal	0.072	-	0.064	-	-	-	-	-
A6-95	Normal	0.125	-	0.159	-	-	-	-	-
A6-96	Normal	0.121	-	0.12	-	-	-	-	-
Cut-off		0.284		0.266					

One of skill in the art will appreciate that the order of the individual antigens within the fusion protein may be changed and that comparable activity would be expected provided each of the epitopes is still functionally available. In addition, truncated forms of the proteins containing active epitopes may be used in the  
5 construction of fusion proteins.

From the foregoing, it will be appreciated that, although specific embodiments of the invention have been described herein for the purpose of illustration, various modifications may be made without deviating from the spirit and scope of the  
10 invention.

## CLAIMS

We claim:

1. A polypeptide comprising an antigenic portion of a soluble *M. tuberculosis* antigen, or a variant of said antigen that differs only in conservative substitutions and/or modifications, wherein said antigen has an N-terminal sequence selected from the group consisting of:

- (a) Asp-Pro-Val-Asp-Ala-Val-Ile-Asn-Thr-Thr-Cys-Asn-Tyr-Gly-Gln-Val-Val-Ala-Ala-Leu (SEQ ID NO: 115);
- (b) Ala-Val-Glu-Ser-Gly-Met-Leu-Ala-Leu-Gly-Thr-Pro-Ala-Pro-Ser (SEQ ID NO: 116);
- (c) Ala-Ala-Met-Lys-Pro-Arg-Thr-Gly-Asp-Gly-Pro-Leu-Glu-Ala-Ala-Lys-Glu-Gly-Arg (SEQ ID NO: 117);
- (d) Tyr-Tyr-Trp-Cys-Pro-Gly-Gln-Pro-Phe-Asp-Pro-Ala-Trp-Gly-Pro (SEQ ID NO: 118);
- (e) Asp-Ile-Gly-Ser-Glu-Ser-Thr-Glu-Asp-Gln-Gln-Xaa-Ala-Val (SEQ ID NO: 119);
- (f) Ala-Glu-Glu-Ser-Ile-Ser-Thr-Xaa-Glu-Xaa-Ile-Val-Pro (SEQ ID NO: 120);
- (g) Asp-Pro-Glu-Pro-Ala-Pro-Pro-Val-Pro-Thr-Thr-Ala-Ala-Ser-Pro-Pro-Ser (SEQ ID NO: 121);
- (h) Ala-Pro-Lys-Thr-Tyr-Xaa-Glu-Glu-Leu-Lys-Gly-Thr-Asp-Thr-Gly (SEQ ID NO: 122);
- (i) Asp-Pro-Ala-Ser-Ala-Pro-Asp-Val-Pro-Thr-Ala-Ala-Gln-Leu-Thr-Ser-Leu-Leu-Asn-Ser-Leu-Ala-Asp-Pro-Asn-Val-Ser-Phe-Ala-Asn (SEQ ID NO: 123); and
- (j) Ala-Pro-Glu-Ser-Gly-Ala-Gly-Leu-Gly-Gly-Thr-Val-Gln-Ala-Gly; (SEQ ID NO: 131)

wherein Xaa may be any amino acid.

2. A polypeptide comprising an immunogenic portion of an *M. tuberculosis* antigen, or a variant of said antigen that differs only in conservative substitutions and/or modifications, wherein said antigen has an N-terminal sequence selected from the group consisting of:
  - (a) Asp-Pro-Pro-Asp-Pro-His-Gln-Xaa-Asp-Met-Thr-Lys-Gly-Tyr-Tyr-Pro-Gly-Gly-Arg-Arg-Xaa-Phe; (SEQ ID NO: 124) and
  - (b) Xaa-Tyr-Ile-Ala-Tyr-Xaa-Thr-Thr-Ala-Gly-Ile-Val-Pro-Gly-Lys-Ile-Asn-Val-His-Leu-Val; (SEQ ID NO: 132), wherein Xaa may be any amino acid.
3. A polypeptide comprising an antigenic portion of a soluble *M. tuberculosis* antigen, or a variant of said antigen that differs only in conservative substitutions and/or modifications, wherein said antigen comprises an amino acid sequence encoded by a DNA sequence selected from the group consisting of the sequences recited in SEQ ID NOS: 1, 2, 4-10, 13-25, 52, 94 and 96, the complements of said sequences, and DNA sequences that hybridize to a sequence recited in SEQ ID NOS: 1, 2, 4-10, 13-25, 52, 94 and 96 or a complement thereof under moderately stringent conditions.
4. A polypeptide comprising an antigenic portion of a *M. tuberculosis* antigen, or a variant of said antigen that differs only in conservative substitutions and/or modifications, wherein said antigen comprises an amino acid sequence encoded by a DNA sequence selected from the group consisting of the sequences recited in SEQ ID NOS: 26-51, 133, 134, 158-178, 196, 235, 237-242, 248-251, 290-293, 304, 311, 313-315, 317, 319, 323, 324, 328, 330, 332, 334 and 336, the complements of said sequences, and DNA sequences that hybridize to a sequence recited in SEQ ID NOS: 26-51, 133, 134, 158-178, 196, 235, 237-242, 248-251, 290-293, 304, 311, 313-315, 317, 319, 323, 324, 328, 330, 332, 334 and 336, or a complement thereof under moderately stringent conditions.
5. A DNA molecule comprising a nucleotide sequence encoding a polypeptide according to any one of claims 1-4.



6. A recombinant expression vector comprising a DNA molecule according to claim 5.
7. A host cell transformed with an expression vector according to claim 6.
8. The host cell of claim 7 wherein the host cell is selected from the group consisting of *E. coli*, yeast and mammalian cells.
9. A method for detecting *M. tuberculosis* infection in a biological sample, comprising:
  - (a) contacting a biological sample with one or more polypeptides according to any of claims 1-4; and
  - (b) detecting in the sample the presence of antibodies that bind to at least one of the polypeptides, thereby detecting *M. tuberculosis* infection in the biological sample.
10. A method for detecting *M. tuberculosis* infection in a biological sample, comprising:
  - (a) contacting a biological sample with a polypeptide having an N-terminal sequence selected from the group consisting of sequences provided in SEQ ID NO: 129 and 130; and
  - (b) detecting in the sample the presence of antibodies that bind to at least one of the polypeptides, thereby detecting *M. tuberculosis* infection in the biological sample.
11. A method for detecting *M. tuberculosis* infection in a biological sample, comprising:
  - (a) contacting a biological sample with one or more polypeptides encoded by a DNA sequence selected from the group consisting of SEQ ID NOS: 3, 11, 12, 135, 136, 151-155, 184-188, 194-195, 198, 210-220, 232, 234, 256-271, 287, 288, 298-303, 305-310, 312, 316, 318, 320-322, 325-327, 329, 331, 333, 335 and 337, the complements of said

sequences, and DNA sequences that hybridize to a sequence recited in SEQ ID NOS: 3, 11, 12, 135, 136, 151-155, 184-188, 194-195, 198, 210-220, 232, 234, 256-271, 287, 288, 298-303, 305-310, 312, 316, 318, 320-322, 325-327, 329, 331, 333, 335 and 337; and

(b) detecting in the sample the presence of antibodies that bind to at least one of the polypeptides, thereby detecting *M. tuberculosis* infection in the biological sample.

12. The method of any one of claims 9-11 wherein step (a) additionally comprises contacting the biological sample with a 38 kD *M. tuberculosis* antigen and step (b) additionally comprises detecting in the sample the presence of antibodies that bind to the 38 kD *M. tuberculosis* antigen.

13. The method of any one of claims 9-11 wherein the polypeptide(s) are bound to a solid support.

14. The method of claim 13 wherein the solid support comprises nitrocellulose, latex or a plastic material.

15. The method of any one of claims 9-11 wherein the biological sample is selected from the group consisting of whole blood, serum, plasma, saliva, cerebrospinal fluid and urine.

16. The method of claim 15 wherein the biological sample is whole blood or serum.

17. A method for detecting *M. tuberculosis* infection in a biological sample, comprising:

(a) contacting the sample with at least two oligonucleotide primers in a polymerase chain reaction, wherein at least one of the oligonucleotide primers is specific for a DNA molecule according to claim 5; and

(b) detecting in the sample a DNA sequence that amplifies in the presence of the oligonucleotide primers, thereby detecting *M. tuberculosis* infection.

18. The method of claim 17, wherein at least one of the oligonucleotide primers comprises at least about 10 contiguous nucleotides of a DNA molecule according to claim 5.

19. A method for detecting *M. tuberculosis* infection in a biological sample, comprising:

(a) contacting the sample with at least two oligonucleotide primers in a polymerase chain reaction, wherein at least one of the oligonucleotide primers is specific for a DNA sequence selected from the group consisting of SEQ ID NOS: 3, 11, 12, 135, 136, 151-155, 184-188, 194-195, 198, 210-220, 232, 234, 256-271, 287, 288, 298-303, 305-310, 312, 316, 318, 320-322, 325-327, 329, 331, 333, 335 and 337; and

(b) detecting in the sample a DNA sequence that amplifies in the presence of the first and second oligonucleotide primers, thereby detecting *M. tuberculosis* infection.

20. The method of claim 19, wherein at least one of the oligonucleotide primers comprises at least about 10 contiguous nucleotides of a DNA sequence selected from the group consisting of SEQ ID NOS: 3, 11, 12, 135, 136, 151-155, 184-188, 194-195, 198, 210-220, 232, 234, 256-271, 287, 288, 298-303, 305-310, 312, 316, 318, 320-322, 325-327, 329, 331, 333, 335 and 337.

21. The method of claims 17 or 19 wherein the biological sample is selected from the group consisting of whole blood, sputum, serum, plasma, saliva, cerebrospinal fluid and urine.

22. A method for detecting *M. tuberculosis* infection in a biological sample, comprising:

(a) contacting the sample with one or more oligonucleotide probes specific for a DNA molecule according to claim 5; and

(b) detecting in the sample a DNA sequence that hybridizes to the oligonucleotide probe, thereby detecting *M. tuberculosis* infection.

23. The method of claim 22 wherein the probe comprises at least about 15 contiguous nucleotides of a DNA molecule according to claim 5.

24. A method for detecting *M. tuberculosis* infection in a biological sample, comprising:

(a) contacting the sample with one or more oligonucleotide probes specific for a DNA sequence selected from the group consisting of SEQ ID NOS: 3, 11, 12, 135, 136, 151-155, 184-188, 194-195, 198, 210-220, 232, 234, 256-271, 287, 288, 298-303, 305-310, 312, 316, 318, 320-322, 325-327, 329, 331, 333, 335 and 337; and

(b) detecting in the sample a DNA sequence that hybridizes to the oligonucleotide probe, thereby detecting *M. tuberculosis* infection.

25. The method of claim 24 wherein the oligonucleotide probe comprises at least about 15 contiguous nucleotides of a DNA sequence selected from the group consisting of SEQ ID NOS: 3, 11, 12, 135, 136, 151-155, 184-188, 194-195, 198, 210-220, 232, 234, 256-271, 287, 288, 298-303, 305-310, 312, 316, 318, 320-322, 325-327, 329, 331, 333, 335 and 337.

26. The method of claims 22 or 24 wherein the biological sample is selected from the group consisting of whole blood, sputum, serum, plasma, saliva, cerebrospinal fluid and urine.

27. A method for detecting *M. tuberculosis* infection in a biological sample, comprising:

(a) contacting the biological sample with a binding agent which is capable of binding to a polypeptide according to any one of claims 1-4; and

(b) detecting in the sample a protein or polypeptide that binds to the binding agent, thereby detecting *M. tuberculosis* infection in the biological sample.

28. A method for detecting *M. tuberculosis* infection in a biological sample, comprising:

(a) contacting the biological sample with a binding agent which is capable of binding to a polypeptide having an N-terminal sequence selected from the group consisting of sequences provided in SEQ ID NO: 129 and 130; and

(b) detecting in the sample a protein or polypeptide that binds to the binding agent, thereby detecting *M. tuberculosis* infection in the biological sample.

29. A method for detecting *M. tuberculosis* infection in a biological sample, comprising:

(a) contacting the biological sample with a binding agent which is capable of binding to a polypeptide encoded by a DNA sequence selected from the group consisting of SEQ ID NOS: 3, 11, 12, 135, 136, 151-155, 184-188, 194-195, 198, 210-220, 232, 234, 256-271, 287, 288, 298-303, 305-310, 312, 316, 318, 320-322, 325-327, 329, 331, 333, 335 and 337, the complements of said sequences, and DNA sequences that hybridize to a sequence recited in SEQ ID NOS: 3, 11, 12, 135, 136, 151-155, 184-188, 194-195, 198, 210-220, 232, 234, 256-271, 287, 288, 298-303, 305-310, 312, 316, 318, 320-322, 325-327, 329, 331, 333, 335 and 337; and

(b) detecting in the sample a protein or polypeptide that binds to the binding agent, thereby detecting *M. tuberculosis* infection in the biological sample.

30. The method of any one of claims 27-29 wherein the binding agent is a monoclonal antibody.

31. The method of any one of claims 27-29 wherein the binding agent is a polyclonal antibody.
32. A diagnostic kit comprising:
- (a) one or more polypeptides according to any of claims 1-4; and
  - (b) a detection reagent.
33. A diagnostic kit comprising:
- (a) one or more polypeptides having an N-terminal sequence selected from the group consisting of sequences provided in SEQ ID NO: 129 and 130; and
  - (b) a detection reagent.
34. A diagnostic kit comprising:
- (a) one or more polypeptides encoded by a DNA sequence selected from the group consisting of SEQ ID NOS: 3, 11, 12, 135, 136, 151-155, 184-188, 194-195, 198, 210-220, 232, 234, 256-271, 287, 288, 298-303, 305-310, 312, 316, 318, 320-322, 325-327, 329, 331, 333, 335 and 337, the complements of said sequences, and DNA sequences that hybridize to a sequence recited in SEQ ID NOS: 3, 11, 12, 135, 136, 151-155, 184-188, 194-195, 198, 210-220, 232, 234, 256-271, 287, 288, 298-303, 305-310, 312, 316, 318, 320-322, 325-327, 329, 331, 333, 335 and 337; and
  - (b) a detection reagent.
35. The kit of any one of claims 32-34 wherein the polypeptide(s) are immobilized on a solid support.
36. The kit of claim 35 wherein the solid support comprises nitrocellulose, latex or a plastic material.
37. The kit of any one of claims 32-34 wherein the detection reagent comprises a reporter group conjugated to a binding agent.

38. The kit of claim 37 wherein the binding agent is selected from the group consisting of anti-immunoglobulins, Protein G, Protein A and lectins.
39. The kit of claim 37 wherein the reporter group is selected from the group consisting of radioisotopes, fluorescent groups, luminescent groups, enzymes, biotin, dye particles and colloidal particles.
40. A diagnostic kit comprising at least two oligonucleotide primers, at least one of the oligonucleotide primers being specific for a DNA molecule according to claim 5.
41. A diagnostic kit according to claim 40, wherein at least one of the oligonucleotide primers comprises at least about 10 contiguous nucleotide of a DNA molecule according to claim 5.
42. A diagnostic kit comprising a at least two oligonucleotide primers, at least one of the primers being specific for a DNA sequence selected from the group consisting of SEQ ID NOS: 3, 11, 12, 135, 136, 151-155, 184-188, 194-195, 198, 210-220, 232, 234, 256-271, 287, 288, 298-303, 305-310, 312, 316, 318, 320-322, 325-327, 329, 331, 333, 335 and 337.
43. A diagnostic kit according to claim 42, wherein at least one of the oligonucleotide primers comprises at least about 10 contiguous nucleotide of a DNA sequence selected from the group consisting of SEQ ID NOS: 3, 11, 12, 135, 136, 151-155, 184-188, 194-195, 198, 210-220, 232, 234, 256-271, 287, 288, 298-303, 305-310, 312, 316, 318, 320-322, 325-327, 329, 331, 333, 335 and 337.
44. A diagnostic kit comprising at least one oligonucleotide probe, the oligonucleotide probe being specific for a DNA molecule according to claim 5.

45. A kit according to claim 44, wherein the oligonucleotide probe comprises at least about 15 contiguous nucleotides of a DNA molecule according to claim 5.

46. A diagnostic kit comprising at least one oligonucleotide probe, the oligonucleotide probe being specific for a DNA sequence selected from the group consisting of SEQ ID NOS: 3, 11, 12, 135, 136, 151-155, 184-188, 194-195, 198, 210-220, 232, 234, 256-271, 287, 288, 298-303, 305-310, 312, 316, 318, 320-322, 325-327, 329, 331, 333, 335 and 337.

47. A kit according to claim 46, wherein the oligonucleotide probe comprises at least about 15 contiguous nucleotides of a DNA sequence selected from the group consisting of SEQ ID NOS: 3, 11, 12, 135, 136, 151-155, 184-188, 194-195, 198, 210-220, 232, 234, 256-271, 287, 288, 298-303, 305-310, 312, 316, 318, 320-322, 325-327, 329, 331, 333, 335 and 337.

48. A monoclonal antibody that binds to a polypeptide according to any of claims 1-4.

49. A polyclonal antibody that binds to a polypeptide according to any of claims 1-4.

50. A fusion protein comprising two or more polypeptides according to any one of claims 1-4.

51. A fusion protein comprising one or more polypeptides according to any one of claims 1-4 and ESAT-6 (SEQ ID NO: 99).

52. A fusion protein comprising a polypeptide having an N-terminal sequence selected from the group of sequences provided in SEQ ID NOS: 129 and 130.



53. A fusion protein comprising one or more polypeptides according to any one of claims 1-4 and the *M. tuberculosis* antigen 38 kD (SEQ ID NO: 150).

54. A diagnostic kit comprising:

- (a) one or more fusion proteins according to any one of claims 50-53; and
- (b) a detection reagent.

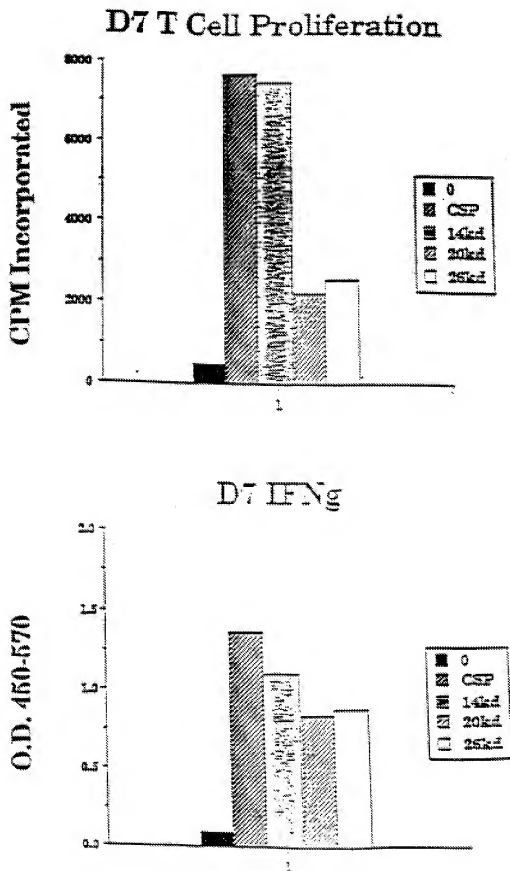
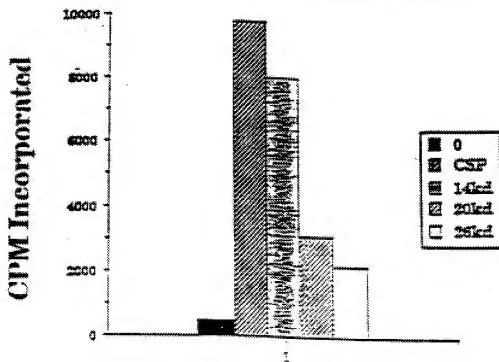


FIG. 14

# D160 T Cell Proliferation



## D160 IFN<sub>g</sub>

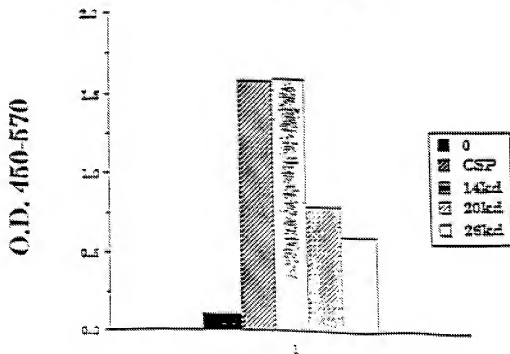
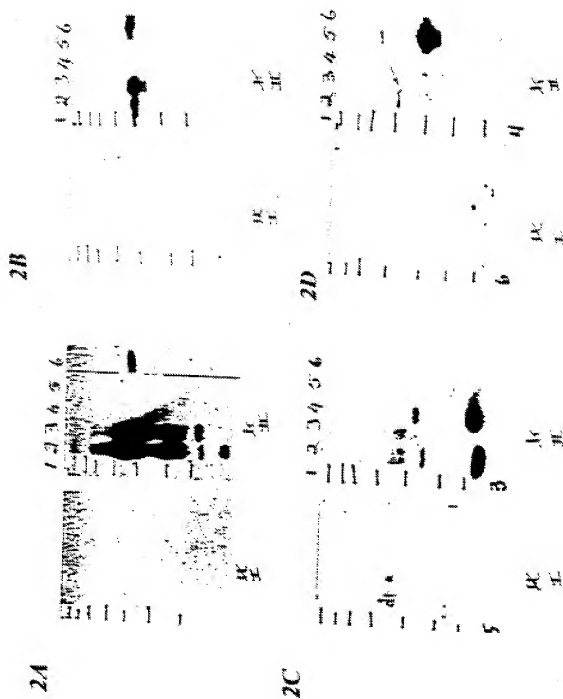


FIG. 1B



FIGS. 2A-D

# T cell clone 131TbH9 responds poorly to CSP

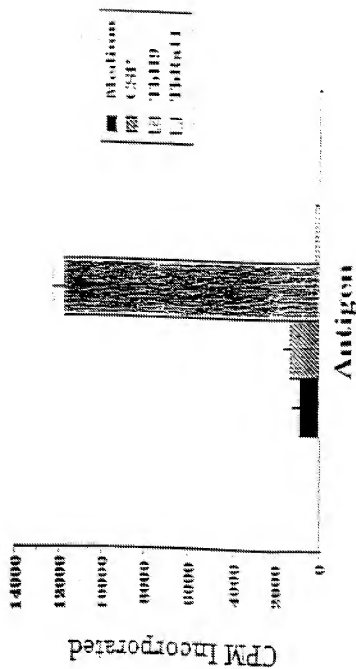


FIG. 3A

# T Cell Clone PPD 800-10 IFN $\gamma$ Production

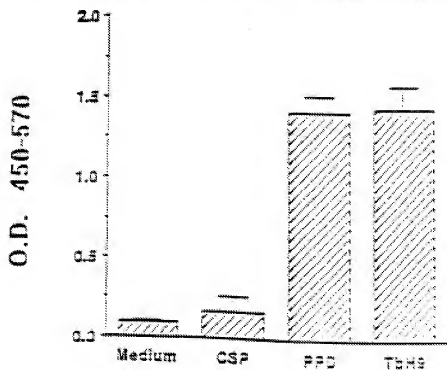


FIG. 3B

Human Anti-Human TGF- $\beta$  Protein Determination

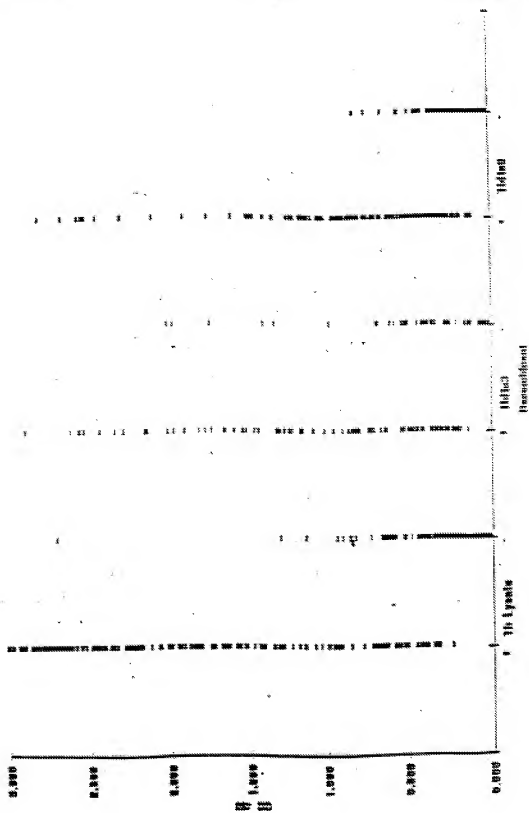


FIG. 4

FIG. 3

ELISA Sensitivity of Hexamethide and Peptides obtained using Human Tip from surrounding membrane

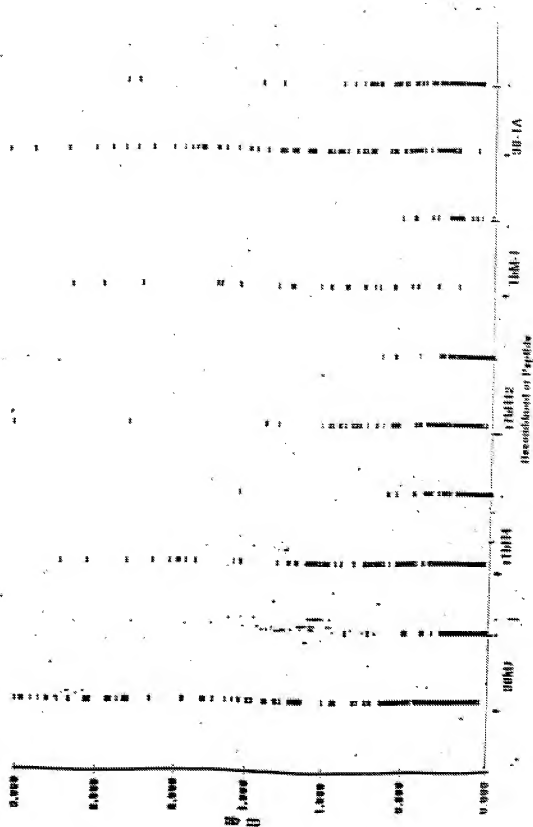




FIG. 6

Activity of Recombinant 30kD and 101kD antigens with sera from *M. tuberculosis* patients, PPD positive and normal donors

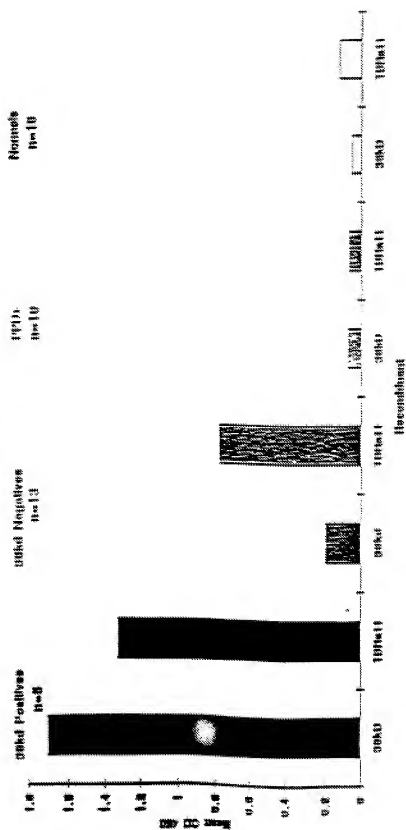
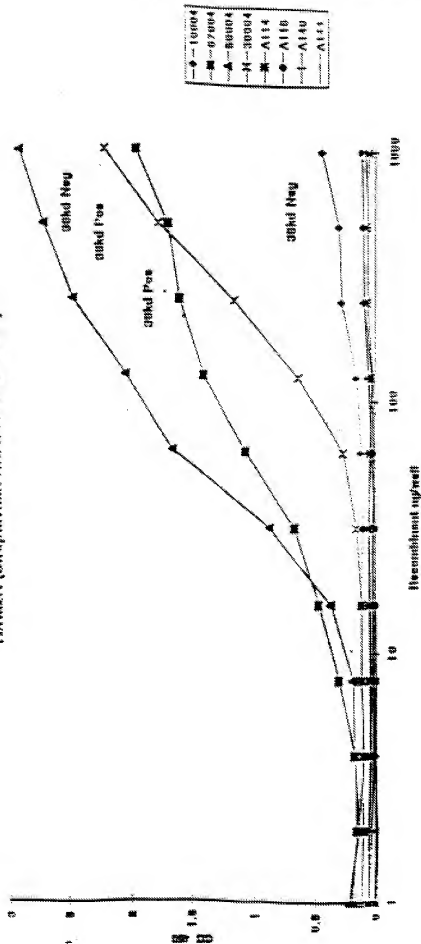


FIG. 7

Tbfta2A (Sheep anti-HRP/Protein A Blot)



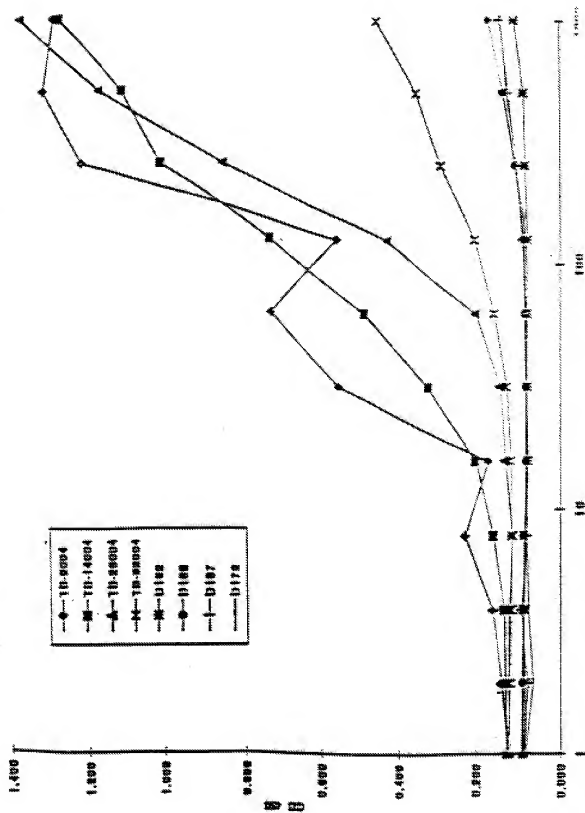
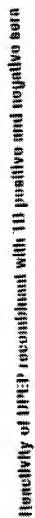
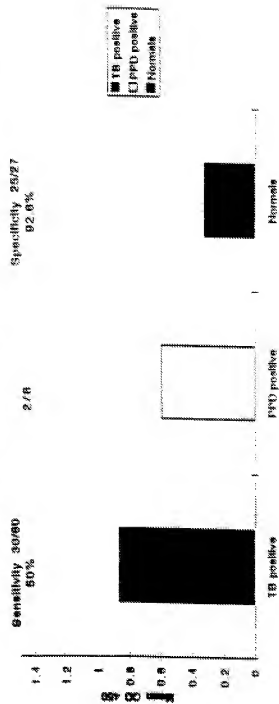
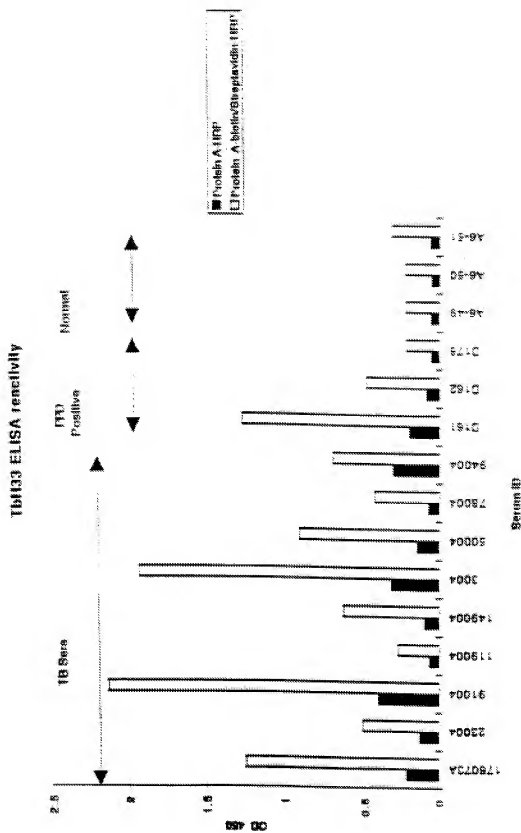


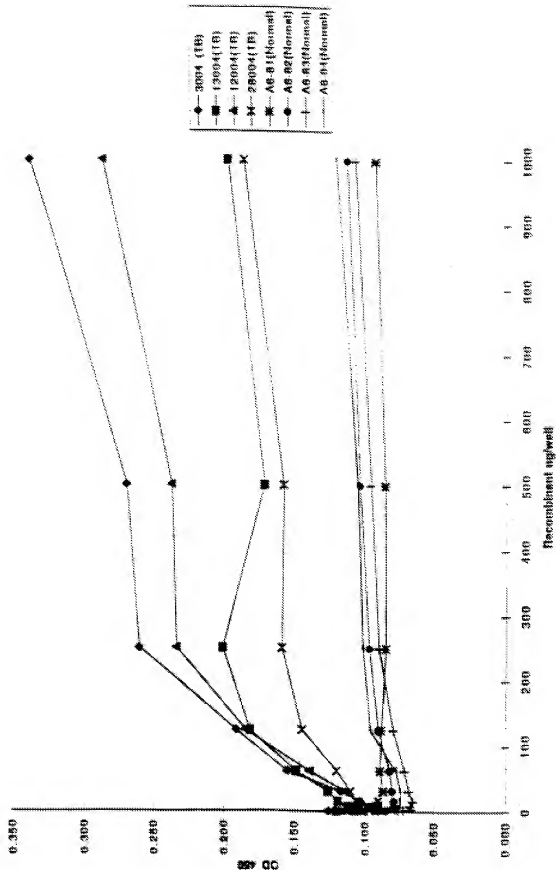
FIG. 9

## TbH2B ELISA reactivity in Streptavidin-IL1P/Protein A-biotin system

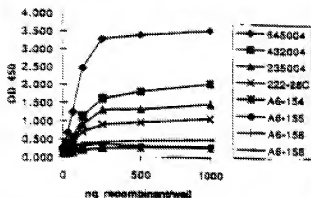


**FIG. 10**

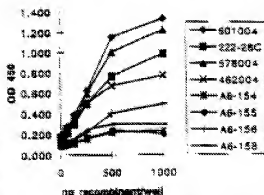
## ELISA reactivity of 1H133



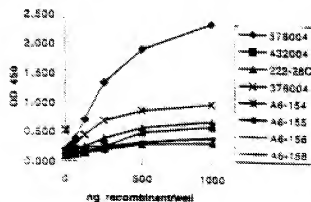
Titration of Mo-1 antigen with TB positive and negative sera



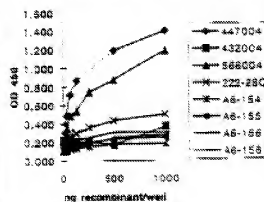
Titration of Mo-2 with TB positive and negative sera



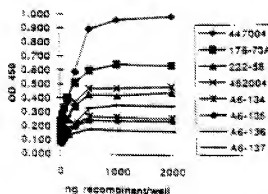
Titration of Mo-4 with TB positive and negative sera



Titration of Mo-28 with TB positive and negative sera



Titration of Mo-29 with TB positive and negative sera



## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

- (i) APPLICANTS: Reed, Steven G.  
Skeiky, Yasir A.W.  
Dillon, Devin C.  
Campos-Neto, Antonia  
Houghton, Raymond  
Vedvick, Thomas S.  
Twardzik, Daniel E.  
Lodes, Michael J.  
Hendrickson, Ronald

(ii) TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF  
TUBERCULOSIS

(iii) NUMBER OF SEQUENCES: 390

## (iv) CORRESPONDENCE ADDRESS:

- (A) ADDRESSEE: SEED and BERRY LLP  
(B) STREET: 6300 Columbia Center, 701 Fifth Avenue  
(C) CITY: Seattle  
(D) STATE: Washington  
(E) COUNTRY: USA  
(F) ZIP: 98104-7092

## (v) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
(D) SOFTWARE: PatentIn Release #1.0, Version #1.10

## (vi) CURRENT APPLICATION DATA:

- (A) APPLICATION NUMBER:  
(B) FILING DATE: 05-MAY-1998  
(C) CLASSIFICATION:

## (viii) ATTORNEY/AGENT INFORMATION:

- (A) NAME: Maul, David J.  
(B) REGISTRATION NUMBER: 31,392  
(C) REFERENCE/DOCKET NUMBER: 210221.417C9

## (ix) TELECOMMUNICATION INFORMATION:

- (A) TELEPHONE: (206) 622-4900  
(B) TELEFAX: (206) 682-6031

## (2) INFORMATION FOR SEQ ID NO:1:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 766 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single



(D) TOPOLOGY: linear

(X1) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CGAGGCACCG GTAGTTTGAA CCAACGCAC AATCGACGGG CAAACGAACG GAAGAACACA	60
ACCATGAAGA TGTGAAATC GATCGCCGCA GGTCTGACCG CGGCGGCTGC AATCGCGGCC	120
GTCGCGGCCG GTGTGACTTC GATCATGGCT GCGGGCCCGG TCGTATACCA GATCGACCGG	180
GTCTCTTCG CGCGCGCACT GCGGTTCGAC CGGGCATCGG CCGCTGACGT CCGGACGGCC	240
GCCCAATTGA CCGAGCTGCT CAACAGCCTC GCGGATCCCA ACCTGTCTTT TCGGAACAAG	300
GGCAGTCTGG TCGAGGCGCG CATCGGGGGC ACCGAGGGCC GCATCGCCGA CCGAAGCTG	360
AAGAAGGCGG CCGAGCACCG GATCTCGCG CTGTGGTTCA GCCTGACGAA CATCCAGCGG	420
GCGCGCGCGG GTTCGGCCAC CCGCGACTTT TCGGTCTCGG CTCGGAAGCT CTCGTGCGCG	480
GTCAAGCAGA AGTCAAGTT CTTGAATCAA GCGCGCTGGA TGCTGTCAGG CGCATCGCGG	540
ATGAGTTTGC TCGAGGCGCG AGGGAACCTG ATTGGCGGCG CGGTTTCAGC CCGCTCTTCA	600
GCTAGCGCGC CGGCTCTGCT ACCCTCTCAT GTCAACACT CCGCGCTGTA GCACGGTGGC	660
GTNTGCGCAG GCGGCGACCG ACCGCGCGGT GCAAGCGGCT CTCGAGATAG GTGTGCTCTC	720
GACAGCAGCG AACACCGCCG NNTGCGGNT TCTGCTGNT GATGA	768

(X2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 768 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(X1) SEQUENCE DESCRIPTION: SEQ ID NO:2:

ATGATCACC ATCACCATCA CGATGAATG ACCTAGAGA CGAGTTCCT GTTCGCGCA	60
GACTTCCTCA CGAGCTTGA CCGTCTCGG CAAGCGGTA CGAGAGCGC GGTCTCCGG	120
GTGGAGGGG TCGCGCGGG CTCGCGCTT CTGATAGTCA AACGAGGCC CAACGCGGG	180
TCCGCTTCC TACTGACCA AGCCATCAG TCGGCTGTC GGCATCCGA CAGCGACATA	240
TTTCTGAGC AGTGTACCG GAGCGCTCG CATCTGAAT TCGGTTGGA AAACAACGAA	300
TTCAATCTG TCGATCTCG GAOTCTCAAG GGCAGTTAG TCAACCGGA GCCCTGGAT	360
TGCGCGCTG TGGGAACCG CGACAGGTC CAGATCGGA AGCTCGGTT GGTGTTCTG	420
ACCGGACCA AGCAAGCGA GATGACCGG AGTACCGGG GCGCTGAGC GCACTCGATA	480

GGCCCGCGCT GGCCGGGATG TCGATCGGGG CGGTCTCTCG ACCTGGCTAG ACCGGATTIT	540
CCCTGATGTC CACCATCTCC AAGATTGGAT TCTGGGAGG CTTGAGGGTC NCGGTGACCC	600
CCCGCCGGGG CTGATPONGG GGTNCGGCGN GGTTCACCC CNTACCNACT GCGNCGGCGN	660
TTGCAATTC NTTCTTCTCT GCGCNAAGG GGACNTTAN CTTGCCGCTN GAAAGGTNA	720
TCCGGGGCCC NTCTNGAAN CCGCTCCCC CT	752

## (2) INFORMATION FOR SEQ ID NO:3:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 813 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (X1) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CATATGCATC ACCATACCA TCACCTCTT AACCGCCAG CGGTCTGGG CGGTCTGACA	60
CCAGCGACAA CGGGCCCGCA TCGATCTGCT AGCTTGAGTC TGCTCAGGCA TGCTCTCAG	120
CAGCCCGATG CCGTATCTTT GTCTCTGACT CAGATATGCG GCGAATCCAA TCTCCCGCT	180
GGCGCGCGCG GTGCTGAAA CTACTCCCGG AGGAATTCG AGCTGCCAT CAAGATCTTC	240
ATCTGCTGCA CGGTCTGCTT TTCTCTCTGT TGTTCGGGTG TGCGCCAGCG CGCGCCCAAG	300
AGCTACTCGG AGGAGTTGAA AGGCACCTAT ACCGGCAGG CGTGCCAGAT TCAATGTCT	360
GACCGCGCTT ACAACATCAA CATCAGCTTG CCGAGTTACT ACCCGGACCA GAAGTCTGTG	420
GAAATTTACA TCGCCGAGAC CGCGGACAAG TTCTCTCAGG CGGCGACATC GTCTACTCCA	480
CGGGAAGCCC CTTAGCAATT GAATATCACC TCGGCCACAT ACCAGTCTCG GATACCGCG	540
CGTGATAGCG AGGCGCTGCT GCTCAGGTG TACCACAAAG CGCGGGGCAC GCACCCAAAG	600
ACCAGCTACA AGCGCTTCCA TTGGGACCGG GCGTATCGCA AGCGAATCAC CTATGACAGG	660
CTGTGCCAGG CTGACACCGA TCGGTGCGCA GTCTCTCTCC CCAITGTTGC AAGGTGAAT	720
GAGCACCGCA GACCGGACA ACNGGTATCG ATAGCGCGCN AATGCCGCTT TGAACCCNG	780
TCAATTTATC ACAACTTCG AGTCACNAA NAA	813

## (2) INFORMATION FOR SEQ ID NO:4:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 447 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(XI) SEQUENCE DESCRIPTION: SEQ ID NO:4:

CGGTATGAAC AGGGCCGCGT CCGATAACTT CGAGCTGTCC CAGGCTGGGC AGCGATTCCG	60
CATTCCGATC GGGCAGGCGA TGGCGATCCG GGGCCAGATC CGATCGGGTG GGGGGTCACC	120
CACCGTTTAT ATCGGGGCTA CCGCTTTCTT CCGCTTGGGT GTTGTGACAA ACAACGGCAA	180
CGCGCGACGA GTCCAAAGCG TGTCGGGAG CCGTCCGGCG GCAAGTCTCG GCATCTCCAC	240
CGCGGACGTG ATCAGCCGCG TCGACGGCGC TCGGATCAAC TCGGCCACCG CGATGGCGGA	300
CGCGTTTAAC GGGCATCATC CCGGTGAGGT CATCTGGGTG AACTGGCAAA CCAAGTCGGG	360
CGGCACCGGT ACAGGGGAAG TGCATTTCG CGAGGGAGCC CCGGCTGTAT TTCGTCTGGG	420
ATACGACCGG CCGGCGCGCC AATTGGA	447

(2) INFORMATION FOR SEQ ID NO:5:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 604 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(XI) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GTCCGACTGC GGTCCGCGAG TATGTGGGTC AGCAAAATGT TGGAGGCCGC CCAACGGGAT	60
CGGTGATCC GAGGTGGCAG TTGTGGAAC CCGCGCGCG GGAAGTATCG GTCCATGCTT	120
AGCCCGCGCA CCGCGAGCGC CGGAATGGCG CGAGTGAGGA GCGCGGCAAT TTGCGGGGCG	180
CGCGGACCG NGAGCGCGCG AATGGCGCGA GTGAGGAGGT GGCAGTCTAT GCCCAAGTGT	240
ATCGAATCAA CCGTGAATCG GACTGCGGCG GCATTGACA ATCGAGTAG TGAACGCAAA	300
TGAATGATGG AAAACGCGCG GAGAGTCTCG ATTTCTGTGT GTGTATAGGT GCTGTGCTGG	360
GTGTGCGGCT ATCAGGATGT TTTTGGCGCA AACTGATGTA CGAGGAACAG GTGTGCGCG	420
NNNNNNNNNN GGGTCTCCNN CCGGNNNTTC TCGNCGANAT CANNAGNCG NTTGATGNGA	480
NAAAAGGGTG GAGCAGNNNN AANTGNGGNN CCAANAANNC NNNNNGNNGG NNAGNTGNTT	540
NNNTNTTNNC ANNNNNNTTG NNNNNNNNNN NNNCAACNN NNNNNNNNA NNGGNTTNTT	600
NAAT	604

(2) INFORMATION FOR SEQ ID NO:6:

- (1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 613 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

TTCCANGTCG AACCACCTCA CTAAAGGGAA CAAAAGCTNG AGCTCCACCG CGGTGCGGGC	60
CGCTCTAGAA CTAGTGGKATM YTYCKGGCTG CAGGAATVCG GYACGASCAT TAGGACAGTC	120
TAACGGTCTT GTTACGGTGA TCGAATGACC GACGACATCC TGCTGATCGA CACCGACGAA	180
CGGTGCGAA CCTTCACCTT CAACCGGGCG CAGTCCCGYA ACGCGCTCTC GCGGCGCGTA	240
CGGATCGGT TTTTCGCGGY GTTGGYCGAC GCGGAGGYCG ACGACGACAT CGACCTCTTC	300
ATCTTGACCG GYCGCGATCC GGTCTTCTGC GCGGACTGCG ACCTCAAGGT AGCTGCGCGG	360
GCAGACCGCG GTGCGGACA TTTCACCGCG GTGGGCGGCC ATGACCAAGC CGTGATCGG	420
CGGATCAAC GCGCGCGCGG TCACGGCGCG GCTCGAACTG GCGCTGTACT CGGACATCTT	480
GATGCGCTCG GAGCACGCGG GCTTCGCGA CACCGACGCG CGGCTGCGGC TGTGCGCCAC	540
CTGCGGACTC AGTGTGTGCT TGCGCGAAA GGTGCGCATC GGNCTGCGGC GGTGATGAG	600
CGTACCGCGC GACTACCTGT CGTGACCGA CGC	613

## (2) INFORMATION FOR SEQ ID NO:7:

## (1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1362 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CGACGAGAC CGCGCGGAG AGCGCGCGCG AACGCGCATC GACCGCGCCC TGCGCAGAT	60
CGGCACTACT CAGGAGGAGG TCGAATCATG AAATTTHICA ACCATATTGA GCGGCTCGGG	120
CGCGCGGAGG CGGCGCGGCG GGTGCGGAG GTCTATGCGG AGCGCGCGCG CGAGTTGCGC	180
CGGTGCGCGG AGCGGCTCGC CATGCTGTCC CGGACGAGG GACTGCTCAC CGCGGCTGCG	240
CGGACGTTGC GCGGACACTT CTGCTGTGCG CAGGTGCGCG GTGCGCGCAA GGAAGCGCTC	300
CGCGCGCGCG TCGCGCGGCG CTGCGGTGCG CGCTGCTGCG TGACGCGACA CACCGACATG	360
CTGTACCGCG CAGGCGCAAC CGACACGCGG GCGCGGATCT TGGCGCGCAC AGCACCTGCG	420
GCGGTGACTC CGAACCGCGG GTATGTGCGG TGGCGCGCAG GAACCGCGAC ACCGCGCGGA	480

CCGCGCGCAC GGTTCGCGCC GGAATGCGCC GCGGAATACC TGGGACCGCC GGTGCAATTC	540
CACATTCATCG CACGCTTGCT CCGTGTGCTG CTGGACGAAA CTTTCTCTCC GGGGGGCGCG	600
CGCGCGCGAC AGCTCAGTCC CCGCGCGCGT GCACTGCTGT TCGCGCGCAA GGTGCGCGCG	660
GAGCATCGCC CGGCGCGCTC CACCGCGCGG CTCGAGCGCC GAAAGCTGCC CGACGATCTG	720
GCATGGGCAA CACCGTCCGA GCGCATAGCA ACCGCGTTCG CGCGCGTCAG CCACGACCTG	780
GACACCGCGC CGACCTGCCC GCGACCGACT CGTCAGGTGG TCAGGCGCGT CGTGGGGTCC	840
TGGCAGCGCG AGCCAAATCC GATGAGCGCT CGCTGAGCGA ACGAAGCACG CGCGGAGCTG	900
CCCGCGGACC TGCACCGGCC CACCGCTCTT GCGCTGCTGA CGCGCGTGGC CGCGGATCAG	960
GTGACCGAGC ACGACGTCGC CGCGCGCGCA TCGCTGCTCG ACACGAGTCC GCGCGTGGTT	1020
GGCGCGTCCG CCGCGCGCGC GTTCACCGCC GCGCGCGCGA TCGCGACCTG GATCGCGCGC	1080
CGCGCGGAGG GCGAGGTGTC GCGCGGAAAAC CGGACTGGGT GAGTGTGCGC GCGCTGTGCG	1140
TAGGCTGTCA TCGCTGCGCC GAGGATGTC GCGCGCGCGA ACGGAGTGGC CGACAGAGGT	1200
GGAAGCTGCG CGGCTGGGT TCGCGCGCAA CGCGCTGCTG GCGCTTGGGT TCGCGGACT	1260
GCGCGATGCG GTCGCGCGCG GCGCTTGGCC GAAGGTCCAG CTCAGCTGCG CGTCACGCAA	1320
GGACCGGACG GTCAAGCGCG GTCAAGCTGC GCGCGGAGG AA	1362

## (1) INFORMATION FOR SEQ ID NO:8:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1458 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (2) SEQUENCE DESCRIPTION: SEQ ID NO:8:

CGGACGACCG GCAATGCTG GCGACCGTAG CGAAAGCGGT CGCGGACCGA CTCGGGCGCG	60
GTATCGCTCC CGTTGAGGAC ATTCAGGACT CGCTGAGGCG CGGCTGGGG GAAGCGGCTC	120
TGATGAGCTT GCGCGCTGTT TACATCATCT ACCGCGACCG GCGCGCGCGA CTCGGGACCG	180
CTAAGGCGTT GCTCGCGCTG CGGACCGAGT TAAAGCTGAG CTGGCGCGCC GTGACGGTAC	240
TGCGCGAGCG CTATCTGCTG CACGACGAGC AGCGCGCGCC GCGCGAGTCC ACCGCGGAGC	300
TGATGAGCTG ATCGCGCGCG TGTGTGCGCG CGCGCGAGGA CCAGTATGAG CGGCGCTCGT	360
CGAGCGGCTG GCGCGAGCGG TTGCGCACCG TATTACGCAA CTGGAATTC CTGCGGAATT	420
CGCGGAGCTT GATGAATCTT GCGACCGAGC TGGAGCTGCT CGCGCGCTGT TTGTTCTTGC	480

CGATTGAGCA TTGCTGCAA TCGATCTTTC	CGAGCTGGG ACAGGCCGCG GAGCTGACG	540
GGCTGAGAG CGGCACCGGA TATCGTTCA	GGCAGCTGCG ACCCGCCGGG GATCGGTTG	550
CGTCCACGG CGGCACGGC AGCGACCGG	TGTCGTTTCT ACCGCTGTAT GACAGTCCG	560
CGGCTGTGCT CTCATGAGC GGTGCGCGC	GTGGCGCTG TATGCTGTG CTGATGTGT	570
CGCACCGGA TATCTGTGAT TTCTCAAGC	CGAAGCGGA ATCCCGCAGC GAGCTCCCG	580
ATTTCACCT ATCGTGTGCT GTGACGAGC	CGTTCTGCG GCGCGTCGAA CGCAACGCG	590
TACACCGCT GGTCAATCG CGAACCGCA	AGATCGTCC GCGGATGCC GCGCGGAGC	600
TGTTGACGC CATCTGCAA GCGCGCAGC	CGGCTGCGA TCGCGGCTG GTGTTTCTG	610
ACAGATGCA TAGCGCAAAC CGGTCGCGG	GGAGAGCGG CATCGAGCG ACCAACGCT	620
GCGCGAGGT CCCACTGCT CGTTACGAGT	CATGTAATC CGGCTCGAT AACCTCGCG	630
GGATGCTCG CGACGCTCG GTGACTGGG	ACCGCTCGA GGAGTCTCG GGTGTGCGG	640
TGCGTTCTT TGATGAGTC ATGATGCA	GCGGCTACC CTTCGCGAA CTGGGTGAG	650
CGCGCGCGC CAGCTGCAAT ATCGGCTCG	GAGTCAAGG TTTCGCGAA CTGCTGCGG	660
CAGTGGGAT TCGGTACGAC AOTGAAGAAG	CGGTCGCTT ACCCAGCGG CTGATGCTC	670
GCATACGCA GCGCGCGAC ACGCATCGC	GGAGGCTGG CGAGAGCGG GCGCATTC	680
CGCGCTTAC CGATAGCGG TTGCGCGCT	CGGCGCGAG GCGCAACGA GAGTCACT	690
CGCTGCTTC GACGCGCA		700

## (2) INFORMATION FOR SEQ ID NO:9:

- (1) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 662 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:9:

ACGGTGTAT CGTCTGGAT CTGGAACCG	GTGGCCGCT ACCTACGAG ATCTACTGG	60
GCGCAGGGG GCTGCGGCTG GGCATCGCG	TCTCTAGT CGGATCGCG GTGGCATCG	120
TCATCGGCTT CGTGCACAGC AGCGCGGTG	CCAAACGGT CAGCGCGAG AACCGGCT	180
GCGCCAGAG CATCTCGGC TCAGCGGAC	CCCAAGCAC CAGCGCGGC GCGCAACCG	240
AAGGTACGC CGCGCGGCG CGCGCGCAG	GCCAAACCT CGAGACAGC AGCGCACCG	300

CCGCGGTGCA GCGCGCGCGG STGCTCAAGG AAGGGGACGA TTGCCCCGAT TGAAGCTGG	360
CGGTCAAAGG TTGACCAAC GCGCGCGAGT ACTACGTGG CGACCAAGCG AAGTTCACCA	420
TGGTGTGCAC CAACATCGCG CTGGTGTCTT GTAAACCGGA GGTTCGGGCC GCGGTGTGGG	480
GCGCCTACGT TTACTGCTG GACAACAAGT GGTGTGTGTC CAACCTGGAC TCGCGCGCCT	540
CGAATGAGAC GCTGTGCAAG AGTTTTTCG CGGCTGAGCA GGTAAAGGAC GCGGTGACCT	600
GGACCTGGAT GGGATCGGCG CGCGCGTGGC CATTGCGCGG GCGCGCGGAT GGGCGCGGCA	660
CGTACAATCT GGTGTGACAA CTGGGCAATC TCGGTGTCTT GCGGTTCGG TTGATCTGA	720
ATCAGCGCGC GCGCGCGCGG GGGCGGTAC CGGCTCGGG TCGAGCGCAG GCGCTCGCG	780
CGGAGTCTCC CGCGCAAGCG GGTAAATTAT TGATCGCTGA TGGTCAATC CGCCAGCTGT	840
GACAACCGCT CGGCTGTGCG CG	862

## (I) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 862 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

TTGATCAGCA GCGCGAAGGT GTACATGCGG TCGGTGGGTG TCGAGGTGAC CAATGACAAA	60
GACACCGCGG GCGCGAAGAT GTGCAAGTA GTGCGCGGTG GTGTGCGCGT GAACGCTGGA	120
GTGCGGAAGG GCGTGTGTGT CACCAAGGTC GACGACCGCG CGATCAACAG CCGCGACGCG	180
TTGTGTGCGG CGGTGCGGTG CAAAGCGCGG GCGCGCACGG TGGCGCTAAC GTTTCAGGAT	240
CGCTCGGCGG GTAGCGCGAC AGTGCAAGTC AGCTCGGCGA AGCGCGAGCA GTGATGAAGG	300
TGCGCGCGCA GTTTCGAAG CTGCGATATA CGGTGGCACC CATGGAACAG GTTTCGGAGT	360
TGGTGTGTGG CGCGCAATTT GTCTGTCTCG TCGACGATCG CACGCGCGAC GCGATGAAG	420
ACCAACCGCG GCGCGTGTTC ACCGAGCTGG TCACCGAGGG CGGTTTGTGT GTGCGAGCGG	480
TGGTGGGGGT GTGCGCGGAC GAGGTGAGAG TCGAAATGCG GCTGAACACA GCGGTGATCG	540
GCGCGGTGGA CGGTGTGCGT TCGGTGCGCG GACCGCGGT GACGCTCTCG GATGTCACCG	600
CGGAAGCCAC CGGAGACATT CT	622

## (II) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1200 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(11) SEQUENCE DESCRIPTION: SEQ ID NO:11:

GGGCGAGCGG TAAGCTGTGT GCGCGCGCGG ACAGTGTGTG TGACAGCATG GCGCGGTGGG	60
ACCAACAGCT CGTCGTGAGG CGCAGGCGGA AGCTGTGGGT CGGTGCATG CCGCGGCAAG	120
AAGGAGCTCC ACTCCAGCGG CTGACCGGCA CAGGAAATG CCATGGACCA GTTCGTCTAT	180
GCCTACGTGC GATCGTGCCC GGGCTACAGG TTGGACTACA AGCGCAACGG GTCCGGTGCC	240
GGCGTGACCC AGTTTCTCAA CAAGGAAACC GATTTCGCGG GCTCGGATG GCCCTTCAAT	300
CGGTCGACCG GTCAACTGTA CCGGTGCGCG GAGCGGTGCG GTTCCCGCGC ATGGGACCTG	360
CGACCGGTGT TCGGCCCGAT CGCGATCACC TACAATATCA AGGCGCTGAG CACGCTGAAT	420
CTTGACCGAC CCACCTACCG CAAGATTTC AACGGCACCA TCACCGTGTG GAATGATCCA	480
CAGATCGAAG CGCTCAACTC CGGACCGGAC CTGCCCGCAA CACCGATTAG GTTATCTTC	540
CGCGACGACA AGTCCCGTAC GTCCGACAACT TTCCAGAAAT AGCTCGACGG TGTATCCAAC	600
GGGCGGTGGG GCAAAGGCGC CAGCGAAACG TTCAGCGGGG GCGTCGCGGT CGGCGCCAGC	660
GCGAACAAAG GAACCTCGGC CTTACTGCAG ACGACCGAGC GGTGGATGAC CTACAGCGAG	720
TGCTCGTTTG CGTGGGTAA GCAATTGAAC ATGGCCGAGA TCATCACGTC GCGCGGTCCG	780
GATCCAGTGG CGATCAGTAC CGAGTCGATC GGTAAAGACA TCGCCCGGGC CAAGATCATG	840
GGACAGGCA ACGACTTGGT ATTGGACAGC TGCTGTTTCT ACAGACCCAC CGAGCTTGGC	900
TCTTACCGGA TGCTGCTGGC GACCTATGAG ATCGTCTGCT CGAAATACCC GGATCGGAGC	960
ACCGGTACTG CGGTAAAGGG GTTTATGCAA GCGCGCATTC GTCCAGGCGA ARAAGGCGTG	1020
GACCAATAGC GCTTCATTCC GTTCCGCAAA TGCTTCAAGC CAAACTTGGC GCGCGCGGTC	1080
AATGCTATT TTGACCTTAG TGAAGGGAAT TCAGCGGTGA GCGATGCCCT TCCGAGGTGA	1140
GGTCCGAAT TTGGGCCGTA TCAGCTATTG CGGCTGCTGG GCGGAGGCGG GATGGCGGAG	1200

(12) INFORMATION FOR SEQ ID NO:12:

(1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1155 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear



## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:12:

GCAAGCAGCT GCAGGTGCTG GTGTTGGAAG AACGGGCAT GCCGAAGACC AAACGCACCA	60
AGACCGGGTA CAGCAGGAT GCCGACGGCG TGCACTCGTT GTTCGACAAG ACCGGGCATC	120
CGTTTCTGCA ACATCTGCTG GCGGACCGCG ACGTCAGCGG GCTCAAGGTC ACCGTGAGG	180
GTTTGTCTGA AGCGGTGGCG GCCGACGGCG GCATCGACAC CAGCTTCAAC CAGACGATCG	240
CGCGGACCGG CGCGCTGTCC TGACCCGAAC CCAACCTGCA GAACATCGCT ATCGGCACCG	300
ACCGCGGCGG GCGGATCCGG GACCGCTTCC TGCTCGGGGA CGGTTACGCC GAGTTGATGA	360
CGCGCGACTA CAGCGGATCG GAGATGCGCA TCATCGGCA CTTGTCCGGG GACGAGGGCC	420
TCATCGAGCG GTTCACACG CGGGAAGACC TGTATTGTT CTTCCGCTCC CGGCTGTTCC	480
GTGTGCCCAT CGACGAGCTG ACGCGCGAGT TCGGCGCGCG GGTCAAGGCG ATGTCTTACG	540
GGCTGGTTTA CGGTTGAGC GCGTACGGCG TGTGGCAGCA GTTGAAATC TCCACCGAAG	600
AAGCCAAAGG GCAGATGGAC GCGTATTTCG CCGGATTCGG CGGGGTGCGC GACTACCTTC	660
GCGCGTAGT CGAGCGGGCG GCGAAGGACC GCTACACCTC GAAGTTGCTG GCGCGTTCGC	720
GCTACCTGCG CGAGCTGGAC AGCAGCAACC GTCAAGTCCG GAGGCGCGCC GAGCGGGCGG	780
GCGTGAAGCG GCGGATCGAG GCGAGCGCGG CCGACATCAT CAAGGTGCGC ATGATCGAGG	840
TCGACGAGCG GGTCAACGAG GCACAGCTGG CGTCGCGCAT GGTGCTCGAG GTCCAGSACC	900
AGCTGTGCTT CGAAATCGCC CGGCTGAGC GCGAGCGGCT CGAGCGCGTC GTCTCGACCA	960
AGATGGGCGG CGGTTACCGG CTCGACGTCG CCGTGGAGGT GTCGGTGGCG TACCGCGGCA	1020
GCTGGGACCG GCGCGGCGAC TGAGTCCGCA GCGTCGATCT GCGGCGGGAA TTGGCGCATT	1080
TTTCCGCTCT GAGTTGAGCG TCGCGCGCAT CCGGACCGAG TTTGTCGAGG GTGTACCGCT	1140
CGAGTAGCTT CTGCA	1155

## (2) INFORMATION FOR SEQ ID NO:13:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1771 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:13:

GAGCGCGCTC TGTGTTTGA ACGTTTTCG CGGTGGCAT CGGACCGGC GTTCCGGGGT	60
---------------------------------------------------------------	----

TGGGGCTCG	GGTTGGCGAT	CGTCAACAG	GTGGTGTCA	ACCACGGCG	ATTGCTGGC	120
ATCGAAGACA	CGACCCGAG	CGGCCAGCC	CCTGGAACCT	GAATTACCT	GCTGCTCCC	180
GGCCGTGGA	TGCGGATTC	GCAGTTCC	GGTGCAGCG	CTGGCGCTG	GAGCAGGAC	240
ATCGAGAACT	CTCGGGTTG	GGCGAACCT	ATCTCAGTG	AATCTCAGT	CACGCGGCA	300
ACCTAGTTGT	GCAGTTACT	TGAAAGCCA	CAGCTATGC	AGTCCACCA	TGCGCAAGT	360
GGCCGAGTA	GTGGGCTAG	TACAGGAAG	GCAACCTAG	GACATGACCA	ATCACCCAG	420
GTATTGGCCA	CGCGCGCAG	AGCGGGGAC	CCAGGTTAT	GCTCAGGGC	AGCAGCAAC	480
GTACAGCCAG	CAGTTGACT	GGGTTACCC	ACCGTCCCC	CGCCCGCAG	CAACCGAGTA	540
CGCTCAACC	TACAGGGCT	TGGTGTGAC	CGGGCGGGT	CTGATACCT	GCTGATTCC	600
GACCATGAG	CGCCCTCTG	GGATGGTTG	CGAACGCGT	CGTCAGCCA	TGTTGGCCAT	660
CGCGCGGGT	ACGATAGCG	TGGTGTCCG	CGGCATCGG	GGCGGGCGG	CATCGCTGT	720
CGGGTTCAAC	CGGCACCGG	CGGCGCCAG	CGGGGGCCA	GTGGCTGCCA	GGCGGGCGC	780
AAGCATCCG	CGAGCAACA	TGCGCGCGG	GTGGGTGGA	CAGGTGGCG	CGAAGTGCT	840
GGCCAGTGT	GTGATGTTG	AAACGATCT	GGCGCGCAG	TGGAGGAGG	GCTCGGCAT	900
CATTGTGCT	GGCAGGGGC	TGATCTTGC	CAACAACAC	GTGATCGCG	CGGCGGCCA	960
GCTCCGCTG	GGCAGTCGG	CGCGAAAC	GACGTAACC	TTCTGTGAG	GGCGACCGC	1020
ACCGTTGAG	GTGGTGGGG	CTGACCCAC	CAGTGATAT	GGGTGGTCT	GTGTTGAGG	1080
CGTCTCGGG	CTCACCCGA	TCTCCGTGG	TTCTGCTCG	GACTGAGGG	TGGTCAGCC	1140
GGTCTGGCG	ATGGGTGCG	CGCTCGGTT	GGAGGCGAC	GTGACCGCG	GGATCTGAG	1200
CGCTCTCAAC	CGTCCAGTG	CGACGACCG	CGAGCGCGC	AACCGAACA	CGGTGCTGA	1260
GGCAATTGAG	AACGACCGG	CGATCAACC	CGGTAACTG	GGGGCGCGC	TGGTGAACAT	1320
GAACGCTAA	CTGTCGAG	TCAACTCGG	CATTGCGAG	CTGGCGCGG	ACTCAGCCGA	1380
TGCGCAGAG	GGCTCGATG	GTCTCGTTT	TGCGATTCA	GTGACCCAG	CGAAGCGCAT	1440
CGCCGACGAG	TTGATCAACA	CGGCAAGGC	GTCAATGCC	TGCGTGGTG	TGCAGGTGAC	1500
CAATGACAA	GACACTCGG	GGCGCAAGAT	CGTCAAGTA	GTGGCGGCTG	GCTCTGCGC	1560
GAACGCTGA	GTGCGGAAG	GGTCTGTTGT	CAACCAAGTC	GACGACCGC	CGATCAACAG	1620
CGCGGACGG	TTGGTTGCG	CGTGGCGTC	CAAGCGCGG	GGCGCACCG	TGGCGCTAAC	1680
CTTTCAGAT	CGCTCGGCG	GTAGCTGCAC	AGTCAAACT	ACCGTGGGA	AGGCGGAGCA	1740

GTGATGAAGG TCGCGCGCGA GTGTTCAAG C

1771

(2) INFORMATION FOR SEQ ID NO:14:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1058 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

CTCCACCCCG GTGCGCGCG GTCTAGAACT AGTGGATCCC CCGGGCTGCA GGAATTGCG	60
ACGAGGATCC GAGCTCCGAG GTTGTGGAAC CCGCGCGCG GGAAGTATCG GTCCATGCT	120
AGCGCGCGCA CCGGAGCGCG GGAATGGCG GAGTGAAGCA GCGGGGCAAT TTGCGCGCGC	180
CGCGCGAGCG CGAGCGCGCG AATGCGCGCA GTGAGGAGCG GGGCACTCAT GCGCAGCTG	240
ATCGAATCAA CCGTCATTCG GCGTGGCGCG CCATTGACA ATCGAGGTAG TGAGCGCAA	300
TGAATGATCG AAAACGCGCG GTGACGTCCG CTGTTCTGCT GGTCTAGCT GCGTGCCTGG	360
GTTTGTGCT ATCAGGATGT TCTTGGCGCA AACCTGATGC CGAGGACAGG GGTGTTCGCG	420
TGAGCGCGAG GCGCTCCGAC CCGCGCGCTC TCGCGAGAT CAGCGAGTGG GTTGATCGCA	480
CAAAAGGCTT GACGAGCTG CAGTAGCGCG TCGAACAAC CGGGAAGTC GACAGCTTGC	540
TGGTATTAC CAGTCCGAT GTGACGTCC GCGCAATCC GCTCGCGCA AAGCGCTAT	600
CGACCTACAA CGAGGAGCAG GTGTCCCGCT TTGCGGTACA AGCGGACAA ATCTCGGTGA	660
AATGTTGCA GAGCTGAGC AATTCGCTT CGATTTCGA AGTGCAACT TCGCGCTGC	720
TGATCGTCC GCGTGGGCT AGCGAGCTGC TGTCCGCTG CAGCAACTC CAAGCGCAAG	780
GTACCGAAGT GATAGACGA ATTTCAGCA CCAAAATCAC CGGACCATC CCGCGAGCT	840
CTGTCAAGT GTTGATCTT GCGCGCAAGA GTGCAAGGCT GCGGACCTG TCGATTGCG	900
AGGAGGCTG CGACCATCT GTCCGAGCA GCATCGACCT CGGATCCGG TCGATTGAG	960
TGCGCGACT GAAATGGAAC GAACCGCTCA ACCTCGACTA GCGCGAAGT GCGTGCAGC	1020
GTGTTGCAA ACGCCCTTT GAACGCTGT AACGNAAC	1058

(2) INFORMATION FOR SEQ ID NO:15:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 543 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:15:

GAATTGGCA CGAGAGTGA TGCATCAT CGGACCGAC CCCCATCTCT GGGAACAGGC	60
GGCGCGGAG GCGGTCCAG GCGCGCGGA TAGCCTCGAT GACATCCGC TCCTCGGGT	120
CATTGACAG GACATGGCG TGCACAGCG CGGCAAGATC ACGTACCGCA TCAAGCTGA	180
ATGTCTGTC AAGATGAGC CGGCGCAAC GCGCTAGCAC GCGCGCGCA GCAAGACCGA	240
AAATGACAG GTTTGCGTT GATTCGTGCG ATTTTGTGTC TGCTCGCGA GCGTACCAG	300
GCGCGCGCA GTTCCGCTG CTGCGGTATC CAGGCGTGA TCCTGATTCC GCGCGCCAG	360
CGCGAGTAA TGCTTCGCT CGACCGGAA TCGCGGATCC CGCGGAGAG TGATCGATGA	420
CGTTGCGAG CGCTCGATG CCGCAATTC CGGAGGAAAC GTGTCGCGAG CGCGTATGA	480
AGCGTCCGA GCGCGCGTG CTGACCGGCT CTGCGTGGC CTCAGTGGC GCGAGCGAG	540
GG	542

(2) INFORMATION FOR SEQ ID NO:16:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 913 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:16:

CGGTGCGGC GCGCGCTGC TTGCGCCGAT TGCGCGCTG GCGCATCAGC TTGCGATCGC	50
CACCATCACC GCGTTTGCG CGGCGACCG CGGTGGCGC GCGCGCGCG ATGCCACGC	120
TTGACCTGG CGCGCGCGC CGCATTTGC ATACAGCAG CGCGCGCGG CACCGTTACC	180
CGGTGCGCA CGGTGCGGC CGTTCGCTT TCAGCGCGG GAGCGCGAAT GACCGCGCG	240
CAGCGCGCG CGCGCGCGC TTGCGCGCT TTGCGCGCG CGCGCGCGG CGCGCAATTG	300
CGGAACAGC AMGCACCGT GCGCGCAGC CGCGCGCGT TAACGCGCT CGCGCGCGC	360
GCGCGCGAG CGCGCATAC CGCGTTGCG GTTGGGTGC CGCGCGTAC CGCGCGCGC	420
GTTCGCGCG AATATTCGC GCGCACGCG AGATCGCGC GCGCGCAT TCGCGCGCG	480
CACCGAACA ACAGCGCAAC GTTGGCGCG CGCGCGCGT TTGCGCGAT CACCGCGCAT	540
TCACCGCGC CACCGCGCT AATGTTTAT AACCGGTAC CGCGAGCGC CGCGCTATTG	600
CGCGCGCGC GAGGCGCTG CGCGCGCGC CGCGACCGC CAAGAGCGG CGGTGCGCG	660

CGGCCCCGCG	GGACCCACCG	GTCCCGCCGA	TCCCGCGGT	GGCGCGGTC	CGGCGCCAT	720
TGCTGCTGCT	GAAGCCGTTA	CGCGCGGTC	CGCGCGGTC	GGCGCGGTC	CGTGGCGCG	780
CGGCCCCGCG	GTTCGCGTAC	AGCCACCGCC	CGTGGCGCG	GTTCGCGCA	TTGCGCGAT	840
TGCGCGCGGT	CGCGCGGTC	CGCGCGGTC	CGCGCGGTC	CGCGCGGTC	CGCGCGGTC	900
CGCGCGGCG	CGC					913

## (2) INFORMATION FOR SEQ ID NO:17:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1872 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

GACTACGTTG	GTGTAGAAA	ATCTGCGGC	CGGACCGCT	AAGCTGGGA	CAATTCTGA	60
TAGCTACCGG	GACACAGGAG	GTTACCGGAT	GGCAATTCG	CGCGCGCGCT	CAGTCAGGTG	120
GTGATGTTG	CTGAGCGTC	TGGCTGCGCT	CGGCTGCGGC	CTGCGCGCG	CGCGCGCGCA	180
GGCGCGCGCG	CGGCGCTTGT	CGGACCGCG	GTTCGCGCG	TTGCGCGCG	TGCGCGCGCA	240
CGCGCGCGCG	ATGCTGCGCG	AAGTGGCGCG	ACAGGTGCG	AACATCAACA	CCAAACTGGG	300
CTACAACAAC	GGCGTGGCG	CGGCGCGCG	CATCGTCATC	GATCGCGCG	GTGCTGCGCT	360
GACCAACAAC	CACGTCATCG	CGGCGCGCG	CGACATCAAT	CGCTTCAGCG	TGCGCTGCGG	420
CCAAAGCTAC	GGCTGCGGT	TGCTGCGGT	TGACCGCGCG	CAGGATGCG	CGGCTGCGCA	480
GTGCGCGCGT	AGCGTGGCG	TGCGTGGCG	GGCGATCGGT	GGCGCGCGTC	CGGCTGCGCA	540
GGCGTGGCTG	CGGATGGCG	ACAGCGCGCG	CGGCGCGCG	ACCGCGCGCG	CGGCTGCGCG	600
CAGGCTGGCT	CGCGTGGCG	AAAGCTGGCG	CGCGTGGCG	TGCGTGGCG	GTGCGCGCG	660
GACATTCGAG	GGCTGCGCG	AGCTGCGCG	CGCAATCGAG	CGCGTGGCT	CGGCGCGCG	720
GTGCTGCGCG	GGCTGCGCG	AGCTGCGCG	TATGACAGCG	CGCGTGGCG	ATAACTTCGA	780
GTGCTGCGCG	GGCTGCGCG	GATTCGCGCG	TGCGATCGCG	CAGGCGCGCG	CGATCGCGCG	840
CGCAATTCGA	TGCGTGGCG	GGTACCGCG	GTTTCATATC	GGCGTTCAGG	CCTTCTGCGG	900
GTGCTGCGGT	GTGCGCAACA	ACCGCAACCG	CGGCGCGCG	CAACCGCGCG	TGCGGAAGCG	960
TGCGCGCGCG	AGTCTGCGCG	TGCGCGCGCG	CGGCTGCGCG	ACCGCGCGCG	ACCGCGCGCG	1020

GATCAACTCG GCCACCGCGA TGGCGGACGC GCTTAAAGGG CATCATCCCG GTGACGTGAT	1080
CTGGTGAAC TGGCAACCA AGTCGGGCGG CACGGGTACA GGGAACTGA CATTCGCCGA	1140
GGGACGCGCG GCGTGATTG TCGCGGATAC CACCGCGCGG CGGCGCAATY GCAATTGCGCG	1200
CAGCGGTGAT TCGCGGTGA GCGCGGAGT TCGGTCTCC GTGCGCGTGG CATTTGGAA	1260
GCAATGAAC AGGCAGAACA CAGCGTTGAG CACCGTCCCG TCGAGGGCAG TTACGTCCAA	1320
GGCGGTGTGG TCGAGCATCG GGATGCCAAG GACTTCGGCA GCGCGCGCGC CCGCGCGCGC	1380
GATCGGACCT GGTTAAGCA GCGGTCTTC TACGAGGTGC TGGTCCGGGG GTTCTTCGAC	1440
GCGACGCGCG ACGGTTCCCG CAGTCTGGT GGACTCATCG ATCGCGTGA CTACCTGACG	1500
TGGTTCGCA TCGACTGCAT CTGTTGCGCG CGTTCCTACG ACTCAGCGCT GCGCGACGGC	1560
GTTACGACA TTGCGACTT CTACAAGTG CTGCGCGAAT TCGGCAAGCT GAGGAGTTTC	1620
GTGCGCGTGG TCGACACCGC TCACGCGCGA GGTATCGCA TCATCACCGA CCGTGTGATG	1680
AATCAGACCT CGGAGTGGCA CCGTGTGTT CAGGAGTCCC GCGCGGACCG AGACGAGCGG	1740
TACGCTGACT ATTACGTGTG GAGCGACACT AGCGAGCGCT ACACCGAGCG CGGATCATC	1800
TTCTCGACA CGGAGAGTGC GAATGCTCA TTGATGCTG TCGCGCGACA GTTCTACTG	1860
GCACCGATTC TT	1872

(2) INFORMATION FOR SEQ ID NO:18:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1482 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

CTTCCGCGAA ACCTGATGCC GAGAACACGG GTGTTCCCGT GAGCGCGAGG GGTCCGACC	60
CGCGGTGCT CCGCGAGATC AGGCACTGCG TTGATCGGAC AAAAGGGTTG ACCAGCGTGC	120
ACGTAGCGGT CGGACACACT GCGAAAGTGG ACAGCTTGCT GGTATTATAC AGTCGCGATG	180
TGACGTCCG GCGCAATCGG CTGCGCGCAA AGGCGGTATG GACCTACAAC GACGAGCAGG	240
GTGTCGGGTT TCGGGTACAA GCGCACAACT TCTCGGTGAA ACTGTTCCAC GACTGGAGCA	300
ATGTCGCGTC GATTCTGAA CTGTCAACTT CAGCGTGCT CGATCCTGCG GCTCGGGTGA	360
CGCAGTGCT GTCCGCTGC AGCAACTGCT AAGCGCAAGG TACCGAAGTG ATAGACGGA	420
TTTCGACGAC CAAATCACC GCGACCATCG CCGGAGCTG TGTCAAGATG GTTATCTGT	480

GCGCCAGAG TCACAGGCGG GCGACCGTGT GGATTGCCCA GGACGGCTCG CACCACTCG	540
TCCGAGCGAG CATGACCTC GGAATCGGGT CGATTGAGCT CACCGAGTGG AATGGAACG	500
AACCCCTCAA CCGGACTAG GCGGAAGTGG GGTGACGGG TTGCTGAGAA CCGCCTTGTG	460
AACGGTGTCA ACGGCAACCG AAAACTGACC CCGTGACGGC ATCTGAAAA TGACCCCTTA	720
GACCGGGCGG TTGGTGTTTA TTCTTCGGTG GTTCGGCTG GTGGGACCGG GCGGAGTGG	780
CGGTCTTTGA GCGGGTAGCT GTGGCTTTG AGGGCGACGA CTTCAGCATG GTGGACGAGG	840
CGGTGATCA TGGCGGACG AACGACTCG TCGCCGCGGA AAACGTCGCG CCACCGCGCG	900
AAGGCTTAT TGGAGTGC GATCAAGCTG GCGCGCTCAT ACCCGGAGGA CACCACTGG	960
AAGAAGAGGT TCGCGGCTC GGGCTCAAC GGAATGTAAC CGACTTCGTT AACCAACAGG	1020
AGCGGATAGC GCGCAACCG GGTGAATTCG GGTAGATGC GCGCGGCTG GTGAGCTCG	1080
GCGAACCGTG CTACCATTC GCGGGGCTG GCGAAGACCA CCGATGACC GCGCTGACAC	1140
GCGGTATCG CCGGCGGAC CGCAAGATGA GTCTTCGCG TCGGAGGCG GCGCAAAAA	1200
CACCACTTA TCGCGGCGG TGATGAATC CAGGCTGCC AGATGTGCGA TGTGTGCGG	1260
TTGAGGCGCA CGAGCATGCT CAAAGTGAA CTCTTCCAC GACTTCGAA CCGGAAGCG	1320
GCGGCGCGCG ATCGGCGCT CACCACTG GACTTCGCG GCTGACACTT CCGCTGACG	1380
GCGGCGCGCG AGGATCTCT GGTGCTCCA GTCTCGCGG GCGGCGCAT CCGGAGCGG	1440
GGCACTGAC TCACGAGGG TGGAGCTT CAATGCTTT GT	1482

## (2) INFORMATION FOR SEQ ID NO:19:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 875 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (2) SEQUENCE DESCRIPTION: SEQ ID NO:19:

GAATTCGCA CGACCGCGCG ATAGCTTCTG GCGCGCGCG CACCAAGTGG CTGAGGGTT	60
CGTGTGCGG GCGACCGCG GCGGACCGAC CCGACCGGT GAGGGCTTC AACCGCGGA	120
CGGTACTCG TTGCTGCTGG ACGGACCGAA CCGGCGCGTG GTTGCTTAC ACCCGCGCT	180
CGGTACGAA ATCGGCTACA TCGGGAAGG CGGACTGCGC AGGATGTGG GCGAGAACC	240
GGAGACATG TTCTTCTACA TCACGCTCTA CAACGAGCG TACGTGCGC CCGCGGAGCT	300

GGAGAACTTC GATCCCGAGG GCCTGCTGGG GGGTATCTAC CGTATCACS CGGCCACCGA	360
GCAACGCCACC AACACAGGCG AGATCCTGGC CTCGCGGGTA GCGATGCCCC CGGCGCTGCG	420
GGCAGCAGAG ATGCTGGCGC CCGAGTGGGA TGTGCGCGCC GAGCTGTGGT CGGTGACCAG	480
TTGGGGCGAG CTAACCCCGC ACGGGGTGGT CATCGAGACC GAGAAGCTCC GCCACCCCGA	540
TGCGCCGGCG GGGCTGCCCT ACCTGACGAG AGCGCTGGAG AATGCTCGGG GCCCGGTGAT	600
CGCGGTGTGG GACTGAGTGC GCGCGTCCC CGACGAGATC CGACCGTGGG TGCCCGGCGC	660
ATACCTCAGC TTGGGACCGC ACGGGTTCGG TTTTCCGAC ACTCGGCCCC CGGCTCGTGG	720
TTACTTCAAC ACCCGCGCGC AATCCGAGGT TGGTGGCGGT TTTGGGAGGG GTTGGCGCGG	780
TGACCGGTTG AATATCGACC CATTCGGTGC CGGTCTGGG CCGCCCGCCC AGTTACCGGG	840
ATTCCAGCAA GTTGGGGGGT TGCGCCCGAG TAAGTT	876

## (2) INFORMATION FOR SEQ ID NO:20:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1021 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

ATCCCCCGGG GCTGCAGGAA TTGGGCGCGA GAGACAAAT TCCAGCGCTT AATGCAGGAA	60
CAGATTGATA ACCGATTGAC AGCGGCACAA CAATATGTGG CGATCGCGGT TTATTTGAC	120
AGCGAAGACC TGCTCGCTTT GCGGAAGCAT TTTACAGCC AAGCGGTGGA GGAACGAAAC	180
CATGCAATGA TGCTGTGCA ACACCTGCTC GACGCGGACC TTCGTGTGGA AATTCCCGGC	240
GTAGACAGCG TCGGAAACCA GTTCGACAGA CCGCCCGAGG CACTGGCGCT GCGGCTCGAT	300
CAGGAACGCA CAGTCAGGA CCAAGTGGGT CGGCTGACAG CGTGGGCGG CGACGAGGCG	360
GATTTCTGCG GCGAGCAGTT CATCGAGTGG TTCTTCCAGG AACAGATCGA AGAGGTGGCC	420
TTGATGGCAA CCTGTGTGCG GTTTCGGAT CGGCGCGGGG CCAACCTGTT CGAGCTAGAG	480
AAGTTCTGCG CAGCTGAGT GGAATGTGCG CGGCGCGCAT CAGCGCGCCC GCACCTGCCC	540
GGGGGGCGCG TCTAGATCCC TGCGGGGAT CAGCGAGTGG TCGGTTTGGC CGGCGGTGCT	600
TCCAGCGCAG CCTTGTGTGG GCGGGGTGG TGAGTACCAA TCCAGGCGAC CCGGACCTGC	660
CGGAAAGAT CGATGCTGTC GTACTCATCG ACCTTCCAGG AGTACACGCG CCGGCGCTGA	720
GCTGCCGAGC GGTCAACGAG TTGCGGATAT TCGTTTAAAG CAGGCAGTGA GGTTCGCAAG	780



GGGGTTGGCC GGACCGCCCT GGGCGCACTG CTGCTCAGGT ATCGGGGGGT CTGGCGAGC	840
AACAAAGTGG GCAGGAGGGG TGGAGCCCGC CGGATCGCA GACCGGGGGG GCGAAACGA	900
CATCAACACC GCACGGGATC GATCTGCGGA GGGGGGTGCG GGAATACCGA ACCGGTGTAG	960
GAGCGCCAGC AGTTGTTTTT CCACGACCGA AGGTTTTTGG GGTTCATCGN GCGNHTTAAG	1020
T	1021

(2) INFORMATION FOR SEQ ID NO:21:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 121 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(1) SEQUENCE DESCRIPTION: SEQ ID NO:21:

CGTGGGAGC AACGAAGAA CACAACATG AAGATGCTGA AATGATCGC GCGAGTCTG	60
ACCGCCCGCG CTGCAATCGG CGGCGCTGCG GCGGTGTGA CTGATCAT GCGTGGCGGN	120
CGGTGCTAT ACCAGATGCA GCGGTGCTC TTGCGKCGCG CACTGCCCTT GACCCCGNA	180
TGCGCCCGTG ANGTCCGAC CGGCGCCGAG TGGACGAGC TGCTCAACAG NCTGNCAT	240
CCCAACCTGT CTTTGGGAA CAAGGONAGT CTGCTCGAGG GNGGNATCG NGGNANCAG	300
GGNGNATC GNCNACACA A	321

(2) INFORMATION FOR SEQ ID NO:22:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 173 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(1) SEQUENCE DESCRIPTION: SEQ ID NO:22:

TCTTATGGT TCGGTTGGC GACCGTTTT GGGKCGGGT GTTAAACCG CTGGCCAGC	60
CGATCGAGG GCGCGAGAC CTGCACTCG ATACTCGCG CGGCTGGAG CTCCAGCGC	120
CTTGGTGGT GNACCGGCA GCGCTGAGG AGCGGTGNA GACCGGATC AAGCGGATG	180
ACCGGATGAC CTGATCGGC CCGGKCGAG GCGAGCTGAT GATCGGGAC CGCAAGACCG	240
GCAAAACCG CCGTGTGTG CGGACACAT CCGCAACCA GCGGAAGAA CTGGAGTTC	300
GTGGATCGC AAGAGCAGG TCGCTTTG TATAGTTGG CCGCGGCGA AGAAGCGGA	360

CTTACCATCG CCG

373

## (2) INFORMATION FOR SEQ ID NO:23:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 352 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

GTGACGCCCT GATGGGATTC GTGGGCGGGG CCGGTCCGCT GCGGTGGTG GATCAACAAC	50
TGCTTACCGG GGTGGCGCAA GCTGTGCTGT TGGCTCAGGC AGCGGTGTG CCGGTGGTGT	120
TCTTGACGGC CTGGTACGGG TTGGCGGATT TAGCCGAGAT CAAGGCGGGC GAATCGGTTC	180
TGATCCATCG CGGTACCGCG GTGTGGGCA TGGCGGCTGT GCAGGTGGCT GCGCATGGG	240
GCGTGGAGGT TTCTGTCAAC GCCAGCGGTG GNAAGTGGGA CAGGCTGGGC GCCATGNGGT	300
TTGACGACGA NCCATATCGG NGATTCCGNC ACATNGAAG TTCGANGGA GA	352

## (2) INFORMATION FOR SEQ ID NO:24:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 726 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

GAATCCGGG TTCATTCTGT TCGACGCGG GCTGGCGATA ATCGACGAAG TGATCAAGCC	50
GCGGTTCGGG GCGCTCATGG GTCCACGCGA GTAATCAGCA AGTTCTCTGG TATATCGCAC	120
CTAGCGTCCA GTTCTTTTCC AGATCGCTTT CGTACCGTCA TCGCATGTAC CGGTTGGCGT	180
GCGCGACGCT CATGCTGGGG GCGTGCATCC TGGCGACGGG TGTGGCGGT CTCGGGTTCG	240
GCGCGCAATC CCGACGCCAA ACCGCGCGGG TCGCGACTA CTACTGGTGC CCGGGGCGCG	300
CTTTGACCC CGCATGGGGG CCGAAGTGGG ATCGGTACAG CTGGCATGAC GACTTGCACC	360
GCGACGCGGA GCGCGCGGAC CACAGCCGCT ACTACCGCGG ACCCATCTTC GRAGTTCGGG	420
TGCTTGACGA TCCCGGTGCT GCGCGCGCGC CCGCGGTGTC CCGTGGCGGC GCATAGCGCT	480
CGTTGACCGG GCGGCATCAG CGAATACGGG TATAAACCGG GCGGTGCGGC CCGCAAGCTA	540
CGACTCGCGG CCGGGCAGAT TTACGCTTCC GTGGCGATGG ATCGCGCGGT CCGATGACAG	600
AAATAGGCG ACCGTTTGGG GAACGCGTTC GAGGACGCTT GAAGGGAACG TGTATGAAC	660

GGCGACAGCG CCTCCACCAT CGACATCGAC AAGTTTGTTA GCGGCACACC CGTTGCGCG 720  
ATCGTG 726

## (2) INFORMATION FOR SEQ ID NO:25:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 580 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

GGGACGACG ACCAACCTCG GCGGCACAC CGCCTATGCG TTGATCGAGG CGACCGGGAT 60  
GCTGCGCGAC CATATCCAA GATGCTGGGT GCGGACGTAG CGACCTTTTG ACCAGCGCGG 120  
CTGCGCGATG GCGCGCGCGT GAATCATTC GCGCGGGGCT TGTGCACTTG ATGAACCGGA 180  
ATAGGGAACA ATAGGGGGGT GATTGGCAG TTCAATGTCT GGTATGGCTG GAAATCCAA 240  
GGCGGGGCAT GCTCGGGCGC GACCAAGGTC GCGCAGCGCG GCCAGCGCGA ATCTGGAGGG 300  
AGCACTCAAT GCGCGCGATG AAGCGCGCGA CCGGCGACCG TCGTTTGAA GCAACTAAGG 360  
AGGGGCGCGG CATTGTGATG CGAGTACCAC TTGAGGCTGG CGGTGCGCTG GTCTGCGAGC 420  
TGACACCGGA CGAGCGCGCG GCACTGGGTG ACGACTCGA AGCGCTTACT AGCTAAGACC 480  
AGCGCAACGG CGAATGGTGG GGTTCACCG CACACCTTCG GGTAGATGTC CAGTGTCTCG 540  
TGGCGGATG ATGCCAGGA GAACTGTTGG ATACAGCGCT 580

## (2) INFORMATION FOR SEQ ID NO:26:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 160 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

AACCGAGCGG CCGCGGGTTT TCGCGGGCGG GGGCGGCTCG GCGCGACCG CGCGGCGCGG 60  
GTACCGCGCG GGTGTTTCGG TGTGCGCGCG GCGGCTGGGG CCGGAGGCGA CGGCATCGCG 120  
GCTGTCACGG GTACGTCGGG CAGCACCGCG GGTGATCCG 160

## (2) INFORMATION FOR SEQ ID NO:27:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 372 base pairs

(B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

GACACCGATA CGATGGTAT GTACGCCAAC GTTGTGACA CGCTCGAGGC GTTCACGATC	60
CAGCGCACAC CGACCGGCT GACGATCGCG GATCGGGCCC GTTTCGCGGA GCGCGGTGCC	120
AAGGCGATGG GAATCGACAA GTTGGGGTA ATTATAGCG GAATGGAGCC GTTGTGCT	180
GAACCGAAC AATGGGAGA CGGCAACAC AGTTGGCGT TGGCGCCCG TGTGTTTTC	240
GCCTACGAGC GCACGTACA GACCAAGCC CG	272

(2) INFORMATION FOR SEQ ID NO:28:

(1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 317 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

GCAGCGGCTG GTTCTGGGAC TATCTGGGCA CGTGACGCA GCGCGAGCTG CGCGAGCTGA	60
AGCGGATCGA GCGAGCGGAT CGCTTGCCTG GTTCATGCG CTACCTGGCC GGTATCACC	120
CGCAGGAGCT GAACGTGGCC GAAGCGGCG CGGTATGCG GTTCGAGCG CGGACGATCC	180
GTTCGGATCT GCGTGGCTT GAGACGGTCT ATTCTGTACA TCGCTGCCC GCTTGGTCC	240
GGATCTGAC CGCGAAGATC AAGAAGCGGT CAAAGATCCA CTTCTCCAC ATTGGCTTC	300
CGGCTGGTT GCGCGG	317

(2) INFORMATION FOR SEQ ID NO:29:

(1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 182 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

GATCTGGAG CTGTGATGA ACAGCGTTCC CGGACGCGG CGCGCCAGCA GTTGGGTGA	60
GCAGCGCGG ACCAGCTCG CGGTGGCGAG CATGGTGATG ACCAGTCCG GCTCGGCCAC	120
CGCTTGGCG CGGTACGAA ACACCGGAC AGGTGCGCG CGCGCGCGG AGCGCGGCT	180
CG	182

## (2) INFORMATION FOR SEQ ID NO:30:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 308 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:30:

GATCGCGAAG TTTGGTGAGC AGGTGGTCCA CCGGAAATC TGGGCGCTG CGAGCGGGT	60
CGCGGTTAC GAGGCGAAGA CAGCGCTGTC CGAGCTGCTG CGGCTCGTCT ACGGCGGGCA	120
GAGGTTGAGA TTGCGCGCCS CCGCGAGCCS GTAGCAAGC TTGTGCGCT GCATCTTAT	180
GAGACTCGG GGTTAGGCTT TGACCATGG GTGTACCGC TGCCCGACGA TTGCGAGCT	240
CGGTTGTCAG ACGAGCTGCT CGAACGCTTT CACCGGTGAA GCGCTAGCTC ATCGACACCC	300
ACGTTTGG	308

## (2) INFORMATION FOR SEQ ID NO:31:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 267 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:31:

CGACACGACA GCACTGACG TGATGATGG TCGCGAGCG CATGAGGAC GGAGAGAATC	60
CGGCGGAAGC TCGCGCGCGG CAATGCTCTA TAGTGACGG CGGTAGAGGG CTGCGCGGAT	120
GGCACCGGAC TATTCTGGTG TCGGCTGGC CGGTAAAGC GGTAAAGGA ATGTGAGGGG	180
ACACGATGAG GAATCACACC TACCGAGTGA TCGAGATGCT CGGAGCTGCS CCGGACGCGG	240
TGACGCGGGC AATCCAGGGC GGTCTGG	267

## (2) INFORMATION FOR SEQ ID NO:32:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1539 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:32:

CTGCTCCGGA AAGAAATGGA GGGGACAGGA TGAGCAATCA CAGTACGGA GTGATCGAGA	60
TGGTGGGAG CTGCGCGGAG GGGCTGAGG CGGCAATCCA GGGCGGTCTG GCGCGAGCTG	120

CGCAGACCAT	GCSCCGCTG	GACTGGTTCC	AAGTACAGTC	AATTCGAGGC	CACCTGGTCC	180
ACGGAGCGGT	CGCGCACTTC	CAGGTGACTA	TGAAAGTCGG	CTTCGGCTGG	AGGATTCCTG	240
AACCTTCAAG	CGCGCCGAT	AAGTCAGGTG	CATCATTAA	CGACTTTTCC	AGAACATCCT	300
GACCGCGTGG	AAACCGGTT	CAGCCGACGG	TGGCTGCGCC	GAGCGCGTGC	CTCCMAATC	360
CCTCGGACAA	TTCTCGGGG	CGCGCTACAA	GGAAGTCGGT	GCTGAATTCC	TGGGTATCT	420
GCTCGAGCTG	TGTGGCTGTC	AGCGCGACGA	AGCGGTGCTC	GACGTGCGCT	CGCGCTCGGG	480
CGCGATGGCG	TTCCCGCTCA	CGCGCTATCT	GAACAGCGAG	GGACGCTAGC	CGCGCTTCCA	540
TATCTCGCAG	AAAGCCATCG	CGTGGTGCCG	GGAGCAGATC	ACCTCGGCGC	ACCCCAACTT	600
CGAGTTGAGG	GTCTCGGACA	TCTACAACTC	GCTGTACAAC	CGGAAAGGGA	AATACCAATC	660
ACTAGACTTT	CGCTTTCCAT	ATCCCGATGC	GTCTTGCGAT	GTGGTGTTC	TTACTCGGCT	720
GTTCAACCCG	ATGTTTCCGC	CGGACGTGGA	GCATATCTG	GACGAGATCT	CCCGCGTGGT	780
GAAGCCCGGC	GGAGGATGCC	TGTGCAGCTA	CTTCTTCTCT	AATGACGAGT	CGTTAGCCCA	840
CATCGCGGAA	GGAAAGAGTG	CGCACAACCT	CGACGATGAG	GGACCGGCTT	ATCGGACCAT	900
CGACAAGAAG	CGGCCCGAAG	AAGCAATCGG	CTTGCCGAGG	ACCTTCGTCA	GGGATGTCTA	960
TGGCAAGTTG	GGCTTCGCGG	TGCACGAACC	ATTGCACTAC	GGCTCATGGA	GTGCGCGGGA	1020
ACCACCGCTA	AGCTTTCAGG	ACATCGTCAT	CGCGACCAAA	ACCGCGAGCT	AGGTGCGCAT	1080
CGCGGAAGCA	TGCGGACACT	GTGCGCGGGA	GGCGCGCTGC	CGCGAGGCGG	ATTAGCGCGG	1140
CAGATTAGCC	CGCGCGCGCT	CGCGCGTCCG	AGTACCGCGC	CGCGAATGCG	GTCACCGCGT	1200
GGTAACGACG	CTTGCAGCGC	TGGCGCGCGG	CTTGCAGGAT	CAGGTGCTAG	ATGCGGACAA	1260
AGCGTCGCTG	ATCGGTCTAT	ACCGACGGTG	ACAGCAGCGG	GTGTGCACT	AGCGCGAAGG	1320
CGACCGCGCT	CTCGGGTCTT	GTCCAGCGGA	TGAGCGCGCG	CAAGCGCGCA	TGACCAAGCG	1380
CGCGCATCAC	GTGCGCGGAT	GGCATACCGT	GATAGCGCAAG	ATGAAATTT	AAGGGCACCA	1440
ATAGATTTCG	ATCGCGCAGA	ACTTGCCTGC	GGTTGCGGCT	CAGCGCGCGG	ACGAGCTTCC	1500
GCGACAAGAA	CGGTATGCGG	TGATCTTCGG	CTCTGCGCG			1519

(2) INFORMATION FOR SEQ ID NO:33:

(1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 851 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(X1) SEQUENCE DESCRIPTION: SEQ ID NO:33:

CTCCAGGGTG	CGCTGGATGA	GGGTACCGC	GGGCGAGCC	GAGCTGACCG	CGGCCAGGT	60
CCGGTTTGT	GGGGGGGCT	ACGAGACGGC	GTATGGGCTG	ACGTTGCGCC	CGCCGGTGAT	120
CGCCGAGAAC	CGTGCTGAAC	TGATGATTCT	GATAGCGACC	AACCTCTTGG	GGCAAAACAC	180
CCCGGGGATC	GCCTTCAACG	AGGCCGAATA	CGCGGAGATG	TGGGCCCAAG	ACGCCCGTGC	240
GATGTTTGG	TACGCGCGG	CGACCGCGAC	GGCGACCGCC	ACGTTGCTGC	CGTTTGAAGGA	300
GGCCCGGAG	ATGACCAAGC	CGGTTGGGCT	CCTCGAGCAG	GCCGCCCGGG	TGGAGGAGGC	360
CTCCGACACC	GTCCCGCGCA	ACCGATTGAT	GCACAATGTG	CCCGAGCGCC	TGAAACGATT	420
GGCCGAGGCC	ACGCAAGGCA	CCAGCGCTTC	TTCCAAGCTG	GGTGGCGTGT	GGAAGACGGT	480
CTCGCCGCAT	GGTTCGCGA	TCAACAACAT	GGTGTGATG	GCCAACAACC	ACATGTGAGT	540
GACCAACTCG	GGTGTTGCA	TGACCAACAC	CTGAGCTCG	ATGTTGAAGG	GCTTTGCTCC	600
GGCGCGCGCC	GGCAGGGCG	TGCAAAACGC	GGCGCAAAAC	GGGTCGCGG	CGATGAGCTC	660
GTGCGGAGC	TGCTGGGTT	CTTCGGGTC	GGCGGCTGG	GTGGCGGCA	ACTTGGGTCG	720
GGCGGGCTCG	GTACGGTATG	GTACCGGGA	TGGCGGAAA	TATGCANAAT	CTGGTCGGCG	780
GAACGTTGAT	GGGGGTAAG	GTTTACCGG	GTTTCTGGA	TGGGGTGAAC	TTTGTCAAGG	840
GAACAGTTA						851

(C) INFORMATION FOR SEQ ID NO:34:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 254 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(X1) SEQUENCE DESCRIPTION: SEQ ID NO:34:

GATGATTCGG	GGGGAATTT	GGACGAGATT	CGGCTCGGCG	GATAACCCAA	TCAATGGAAC	60
GTAGATTAT	TGCTTCAGG	GGCCCGAGTA	ATGGTTCGCA	GGAGAGGAAC	CTTACTGCTG	120
CGGCGACGTG	TGTTAGTTC	TGATACCGG	GGAAGGCTC	GACATTTTCC	ACCGACACCC	180
GCATCCAAAC	GTTCGAGGGC	CAGTCAGCT	TGTGAGCGAG	GGGACGAGT	CGCAGGCTGC	240
GCTTGGTCAA	GATC					254

(E) INFORMATION FOR SEQ ID NO:35:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1227 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:15:

GATCTGACC	GAAGCGGCG	CGGCCAAGGC	GAAGTCGCTG	TTGGACCAGG	AGGGACGGGA	60
CGATCTGCG	CTCGGGATCG	CGGTTCAAGC	GCGGGGGTGC	GCTGGATTTC	GCTATAACCT	120
TTTCTTCGAC	GACCGGACGC	TGGATGGTGA	CCAAACCGCG	GAGTTCGGTG	CTGTCAAGTT	180
GATCGTGGAC	CGGATGAGCG	CGCGGTATGT	GGAAAGCGCG	TCGATCGAAT	TGCTCGACAC	240
TATTGAGAA	CAAGGTTTAC	CATCGACAAT	CCCAACGCCA	CCGGTTCCTG	CGCTTGGGG	300
GATTCTGTCA	ACTGATAAAA	CGCTACTACG	ACCCCGCGGT	GCGCAACACG	TACGAGCACA	360
CCAAGACCTG	ACCGCGCTGG	AAAAGCAACT	GAGCGATGCC	TTGCACCTGA	CGCGTGGCG	420
CGCGCGCGCG	CGCAGGTGTC	ACCTGCATGG	TGAACAGCAC	CTGGGCGCTGA	TATTGCGAGC	480
AGTACACGAT	TTGTTCGATC	GAGGTCACTT	CGACCTGGGA	GAAGTGGTTG	CGGAACGGGT	540
CGCTGCTCAG	CTTGGCCAA	GCCTGATCGG	AGCGCTTCTG	GCGCACGCGG	TCGTGGATAC	600
CGCACAGCGC	ATTGCGAAGC	ATGTTGTCCA	CATCGCGGTT	CTCCAGCGCG	TTGAGGTATC	660
CCTGAATCGC	GGTTTTCGCC	GGTCCCTCGG	AGAAATGTGC	TGCCGTGTTG	GCCTCGTTGG	720
TGCGGACCCC	GTATATGATC	GCGCGCTCCA	TAGCGACAC	CAGCGCGAGG	GCTACACAA	780
TGCGGATCGC	CAGCGCGCTG	TGCGTGGCTT	TGGGTAGGA	CACCTGCGCG	GGCAGCGCGG	840
GATATGGCGC	CGGCGGACGC	CGCGCTCTGT	CTGCGGTGC	CGGGCGAAG	GCGGTTCGG	900
CGGCGCGGAG	GTCTTGGGGG	TAGTCCAGGG	CTTGGGGTTC	GTGGGATGAG	GGCTCGGGGT	960
ACGCGCGCGG	TGCTTGGTGT	CGGACCGCG	GCTTGGCGGA	GTGGGACCG	GGCATTGTGG	1020
TTCTCTTAGG	GTGTGGGACG	GGACGAGCTG	CTAGGGCGAC	AACCGCCCGT	CGCTTCAGCG	1080
GGCAGCATCG	GCAATCAAGT	GAGCTTGGTA	GGCAGGCTAG	CGCAACAGCT	GCTTTCAGCT	1140
CTCAACCGGA	CGGGGCGGGG	CGCGGCGCG	ATAATGTTGA	AAGACTAGGC	AACCTTAGGA	1200
ACGAAGGAGC	GAGATTTTGT	GACGATC				1227

## (2) INFORMATION FOR SEQ ID NO:16:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 181 base pairs



(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:36:

GGGTGTTCGG CGGATCCGGC GGGTGGTTGA ACCGCAACGG CGGGCCCGGC GGGGCCGGCG	60
GGACCGCGCG TAACCGTGGT GCGCGCGGCA AGCGCTGGTT GTTCGGGGCC GGCGGGTCCG	120
GGCGCGCGCG CACCAATGGT GCGGTGGCGG GTCCCGCGCG ATTGTCTAC GGCACCGCGG	180
G	181

(2) INFORMATION FOR SEQ ID NO:37:

(1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 290 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:37:

GGGTGTTCGG CGGATCCGGC GGGTGGTTGA ACCGCAACGG CGGTGTTCGG GCGCGCGCGC	60
GGACCGCGCG GTTTCGGGT GCGCGCGGCG AGCGCGGCTT CGGTGGCGAG GCGCGCAATG	120
GCGCGCGCTC CACCGCGCGC AACCGCGCTC TTGGCGGCGC GCGCGGTGGC GGAGCGACCG	180
CGCGCGACCG CGGTTTGGT GGCACCGCGC GTAAGGTGGC CGAGCGCGGN ATTGGCGCGC	240
GCACTCAGAG CGCGACCGCG CTCGCGGTG ACCGCGGTGA CGCGCGTGAC	290

(3) INFORMATION FOR SEQ ID NO:38:

(1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 34 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:38:

GATCCACTGG CATCGGGGT GTCACTGGAA GCAT	34
--------------------------------------	----

(2) INFORMATION FOR SEQ ID NO:39:

(1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 155 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:39:

GATGCTGCTC	CCTCCGCCCC	TTGCCGCCGA	CGGCACCGGT	CCCACCGTGA	CCGAACAGGC	60
TGGCGTGTCT	GCCAGCACCC	CGGCACCGGC	CGACGCCCGA	GTCCACCAAT	GGCACCGTGC	120
TATCCCCACC	ATTGCCCGCG	GNCCTACCGG	CACCG			150

## (2) INFORMATION FOR SEQ ID NO:40:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 53 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ix1) SEQUENCE DESCRIPTION: SEQ ID NO:40:

ATGCGTTCA	CGGGCGCCG	GGACCGGGC	AGCCCGGNG	GGCCCGGGG	TGC	60
-----------	-----------	-----------	-----------	-----------	-----	----

## (2) INFORMATION FOR SEQ ID NO:41:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 132 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ix1) SEQUENCE DESCRIPTION: SEQ ID NO:41:

GATCCACGTC	GGGTCCAGAC	GTTCGCCCGG	GGGCCACCCC	GACCTAGTGC	GGCAACGGCG	60
GCACCGGGCG	CAACGGCGCG	AACGCCACCG	TGCTCGGNG	GGCCCGCGGG	GGCCGGCGCA	120
AGGCCGCCCA	CG					132

## (2) INFORMATION FOR SEQ ID NO:42:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 122 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ix1) SEQUENCE DESCRIPTION: SEQ ID NO:42:

GATCGCGCG	CGGACCGGC	GGGACCGCG	GCAAGCGCG	NAACGGGGCG	CGCGAGCGCA	60
CGGCGCAAGA	ATCTCTCGCG	TCGCGCAATG	GCGCGAATCG	CGGACAGGGC	GGCAACGGCG	120
GCAGCGCGCG	CA					132

## (2) INFORMATION FOR SEQ ID NO:43:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 703 base pairs

(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (X1) SEQUENCE DESCRIPTION: SEQ ID NO:41:

CGGCACGAGG ATCGGTACCC CGCGGCATCG GCAGCTGCCG ATTCGCGCGG TTTCGCCACC	60
CGAGGAAGCG CCTACCGCA TGGCGCTGCC GAAGTAGGCC GATCGGTTCG CGATGCCGGC	120
ATGAACGGCG GCATCAAAAT TAGTGCGGGA ACCTTTCAGT TTAGCGGCGA TAATGCGTAT	180
AGCACTAAGG AGGATGATCC GATATGACGC AGTCCGAGAC CGTGACGGTG GATCAACCAAG	240
AGATTTTGAA CAGGCGCAAC GAGGTGGAGG CCGCGATGCC GGACCCACCG ACTGATGTCC	300
GCATCACACC GTGCGAACTC ACGGCGGNTA AAAACGCCGC CCAACAGNTG GTTNTGTCCG	360
CGGACAACAT CGCGGAATAC CTGGCGCGCG GTGCCAAGGA CGCGCAGCGT CTGGCGAATT	420
CGCTGCGCAA CGCGGCCAAG GGTATGCGCG AGTTTGATGA GGAGGCTGCC ACGCGCTTGG	480
ACAACGACGG CGAAGGAACT GTGCAGGCAG AATCGCGCGG GCGCGTCCGA CGGCGCACTT	540
CGGCGCAACT AACCGATACG CGGAGGCTGG CGACGCGCGG TGAACCCAAAG TTGATGGATC	600
TCAAGAAAGC GCGAAGGAAG CTCGAAACCG CGGACCAAGG CGCATCTCTG GCGCACTGNG	660
GGGATGGGTG GAACACTTNC ACCCTGACGC TGCAAGCCGA CG	702

## (2) INFORMATION FOR SEQ ID NO:44:

(1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 298 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (X1) SEQUENCE DESCRIPTION: SEQ ID NO:44:

GAAGCCCGAC GCTGTCTGG CGACCTGGCG GTCAAAGCGG CATCGCTCGG TGGCGTGGGA	60
GGCGCGCGGG TCGCTTCGCG GCGCTTGGGA TCGCGGATCG GCGCGCGCGA ATCGGTGGCG	120
CGCGCTGGCG CTGCTACACT TCGCGCTTA GCGCAAGGAA GCGCGCGCGG CGCGCGCGCG	180
CTGCGCGCGG GTGCGATGG AATCGCGATG GGTGCGCGCG ATCAGGGACA AGCGGGCGCC	240
AAGTCCAAGG GTTCTCGCA GGAAGAGGAG GCGCTCTACA CGGAGGATCC TCGTGGCG	298

## (2) INFORMATION FOR SEQ ID NO:45:

(1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 1058 base pairs  
(B) TYPE: nucleic acid

(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:45:

CGGCACGAGG ATCGAATCGC GTCCCGGGGA GCACAGCGTC GCACTGCAAC AGTGAGGAG	60
CCATGACCTA CTCGCGGGGT AACCCCGGAT ACCCGCAAGC GCAGCCCGCA GGCTCTACG	120
GAGGGGTGAC ACCCTCGTTC GCCCAAGCCG ATGAGGGTGC GAGCAAGCTA CCGATGTACC	180
TGAACATCGC GGTGGCAGTG CTCGGTCTGG CTGCGTACTT CGCCAGCTTC GGCCTAATGT	240
TCACCTTCAG TACCGAACTC GGGGGGGGTG ATGGCGCAGT GTCCGGTGAC ACTGGGCTGC	300
CGGTGGGGT GGTCTGTGTC GCTCCCTGTC TTCCCGGGGT GGTTCGTGTC CCTAAGGCCA	360
AGAGCCATGT GACGGTAGTT GCGGTGCTCG GGTACTCGG CATTATTCTG ATGGTCTCGG	420
CGACCTTTAA CAAGCCCAAC GCTATTGCA CCGGTTGGGC ATTGTGGGT GTGTGGCTT	480
TCACTGTTT CAGGCGGTT GCGGCAGTCC TGGGGCTCTT GGTGGAGACC GGCCTATCA	540
CGCGCGCGGC GCGCGCGGCC AAGTGGACC GGTATGGACA GTACGGCGCG TACGGGCACT	600
ACGGGAGTGA CCGGGTGACG CCGGGTGGGT ACTACGGTCA GCAGGGTCTT CAGCAGCGCG	660
CGGACTGCA GTGCGCGCGC CGGCAGCACT CTCGCCAGCC TCCCGATAT GGCTCGCACT	720
ACGCGGGCTA TTGCTCCACT CCGAGCCAAAT CGGCGCACTG ATACACTGCT CAGCCCGCGG	780
CGGAGCGGCG GCGGCAGTCC GGGTCGCAAC AATCGGACCA GCGCCGATCC AGCCCACTTA	840
CGGGCTTTCC GAGCTTGAGC CCACCACGAC CGGTCACTGC CCGGACGGGG TGCGAGCTTC	900
GTTCGGCTCC AGTCAACTAT TCAAACCCCA GCGGGGGGGA GCAGTCGTTC TCCCGCGGGG	960
CGCGCGCGGT CTAACCGGCG GTTCGGCGGT CCGGTCCGCG GTGTGCGCGA AGAGTGAACA	1020
GGGTCTCAGC AAGCGCGGAC GATCTCTGTC CGGAATTC	1058

(2) INFORMATION FOR SEQ ID NO:46:

(1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 327 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:46:

CGGCACGAGA GACCGATGCC GCTACCGTCG CGCAGGAGGC AGGTAATTC GAGCGGATCT	60
CGGCGGAGCT GAAAACCGAG ATCGACCGAG TGGAGTGAC GCGAGGTTCC TTGCAGGCCC	120

AGTGGCGCGG CGCGCGGGG ACGGCGCCCC AGGCGCGGT GGTGGCTTC CAAGAAGCAG	180
CCAATAAGCA GAAGCAGGAA CTCGACGAGA TCTGACGAA TATTCGTGAG GCCGCGTCC	240
AATACTCGAG GCGCGACGAG GAGCAGCAGC AGGCGCTGTC CTCGCAATG GCGTCTGAC	300
CGGCTAATAC GAAAAGAAAC GGAGCAA	327

## (2) INFORMATION FOR SEQ ID NO:47:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 170 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (X1) SEQUENCE DESCRIPTION: SEQ ID NO:47:

CGTGGCGGAT GATGCGGTG TCGAACGTGA CGATTCTGT ACCGCGGTG TTGAGATCAA	60
CCAACAACGT GTGGCGTGC CCAATGTGC CGNACCGTG GATCTGGTG ATCTTTTCT	120
TCTTCATCAG GAAGTCGACA CGGCGCACCC TGCCCTCGGN TACCTTTGG	170

## (2) INFORMATION FOR SEQ ID NO:48:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 127 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (X1) SEQUENCE DESCRIPTION: SEQ ID NO:48:

GATCGGGCGG CACGCGGGT GCGGCGGCA GACCGCTGG CGTGGCGGC AACGCGGGG	60
CGGCGGTGG CGGCGAACC GGTGGGTGC TCTTCGCAA CGCGGTGCC GCGGGCAGG	120
GGGCGGT	127

## (2) INFORMATION FOR SEQ ID NO:49:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 81 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (X1) SEQUENCE DESCRIPTION: SEQ ID NO:49:

CGGCGGCAAG GCGGCGACC CGGCGAAGG GAGCGGGCG GCGGCGGCA AGGCGGCAA	60
CGGCGGCTG GCGCTGACC G	81

## (2) INFORMATION FOR SEQ ID NO:50:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 149 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:50:

GATCAGGGCT GGCCTGGCTC GCCCAGAAAG GCGGTAAAGG AGGAGCTGCC GGAATGTTTG	60
GCAACGGCGG GCGCGGGGGT GCGGGGCGGT CCAACCAAGC CGGTAACGGC GGNCGCGGCG	120
GAAACGGTGG TGCCGGGTGG CTGATCTGG	149

## (2) INFORMATION FOR SEQ ID NO:51:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 165 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:51:

CGGCACGAGA TCACACCTAC CGAGTGATCG AGATGTCGG GACCTGCCCC GACGGTGTCC	60
ACCGCGNAAT CCAGGGCGGT CTGGCCCCAG CTGGCCAGAC CATCGCGGCG CTGGAGTGT	120
TGGAATACA GTCAATTGGA GCCACCTGG TCGACGGAGC GGTGCGGCAC TTCCAGGTGA	180
CTATGAAAGT CGGCTTCGCG CTGGAGGATT CCTGAACCTT CAAGCGCGGC CGATAACTGA	240
GGTGATCAT TAAGGACCTT TTCCAGACAA TGTGACCGCG CTGAAACGCG GGTTCGCGCG	300
ACGTTGGCTC CGCCGAGGCG CTGGCTCCAA AATGCTCGCG ACAATTGCTC GCGCG	355

## (2) INFORMATION FOR SEQ ID NO:52:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 999 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:52:

ATGCATCACC ATCACCATCA CATGCATCAG GTGACCCCCA ACTTGACAGC TGCAAGGGA	60
CGATTGCGCG CACTGGCTAT CGGGCGGATG GCGAGCGCCA GCTTGTGACG GATTGCGGTG	120
CGGGCGAGTC CCAACGCGGA TCGGAGGCGA GCGCCCGCGC TACCGACACG GCGCGCTCG	180
CGCGGTGGA CGGCTGACAG GCGACGCGCA CGGCGACGCT CTGTTGCGCG GCGACGACG	240